

**WEST****End of Result Set**

Generate Collection

Print

L3: Entry 2 of 2

File: USPT

Jun 5, 1979

DOCUMENT-IDENTIFIER: US 4157390 A

TITLE: Process for vaccine preparation

**CLAIMS:**

1. A process for the preparation of a vaccine for parenteral administration to pregnant mammals as a means for reducing the incidence of neonatal diarrhoea, in which process K88a,b/adhesion factor of an enteropathogenic strain of E. coli is adsorbed from a cell-free aqueous solution onto erythrocytes possessing their natural surface characteristics, the ratio of adsorbed K88a,b/adhesion factor to the volume of erythrocytes present being in the range 125 to 800 hemagglutination units per ml, and the composition so formed is rendered storable by the addition thereto of a preservative such as sodium azide or formalin.



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:45:22 ; Search time 41.42 seconds  
(without alignments)  
121.122 Million cell updates/sec

Title: US-09-488-737-2

Perfect score: 157

Sequence: 1 MVNKDVKQTAFGAPVMDNNVITAGPRG 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL\_19:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	52.9	492	10	Q92PK5
2	81	51.6	16	2	Q9RS12
3	78	49.7	505	2	Q9RW19
4	75	47.8	492	10	Q92PK6
5	75	47.8	492	10	Q9SYT7
6	75	47.8	492	10	Q9M4X3
7	75	47.8	492	10	Q940G2
8	75	47.8	496	10	Q92PK7
9	74	47.1	41	10	O04821
10	74	47.1	455	2	Q9LA52
11	74	47.1	484	2	Q9RG14
12	74	47.1	505	2	Q9LA51
13	74	47.1	507	16	Q9SUE2
14	73	46.5	492	10	Q9SDS9
15	72	45.9	492	10	O09484
16	72	45.9	493	10	O22472

17	71	45.2	483	2	Q9AQ09	Q9aqg9 bacillus su
18	71	45.2	484	2	P77939	P77939 rhizobium s
19	71	45.2	493	2	Q9EV50	Q9ev50 staphylococ
20	70	44.6	491	10	O22640	O22640 raphanus sa
21	70	44.6	492	10	Q9SB22	Q9sb22 oryza sativ
22	69	43.9	487	2	Q9RJK9	Q9rjk9 streptomyce
23	69	43.9	492	10	Q9ZWH7	Q9zwh7 oryza sativ
24	69	43.9	492	10	Q9ZRI9	Q9zri9 oryza sativ
25	68	43.3	492	10	Q9FEK5	Q9fek5 hevea bras
26	68	43.3	1013	10	Q9LDS9	Q9lds9 arabidopsis
27	67	42.7	492	10	Q9SW99	Q9sw99 manihot esc
28	67	42.7	492	10	Q94EV9	Q94ev9 suaeda mari
29	66	42.0	509	2	Q9FAZ3	Q9faz3 vibrio rumo
30	66	42.0	527	4	Q9BWT9	Q9bwt9 homo sapien
31	65	41.4	527	6	Q9GKV3	Q9gkv3 canis famil
32	64	40.8	493	10	O81337	O81337 mesembryant
33	64	40.8	504	16	Q9JRF5	Q9jrf5 neisseria m
34	64	40.8	526	13	Q918V5	Q918v5 brachydanio
35	62	39.5	506	2	P77924	P77924 pseudomonas
36	61	38.9	484	16	Q9CPK5	Q9cpk5 pasteurella
37	60	38.2	349	16	Q9CI07	Q9ciq7 lactococcus
38	59	37.6	492	10	Q9M504	Q9m504 helianthus
39	59	37.6	492	10	O81336	O81336 mesembryant
40	58	36.9	488	2	P77948	P77948 streptomyce
41	58	36.9	492	10	O24511	O24511 nicotiana t
42	57	36.3	497	5	O18193	O18193 caenorhabdi
43	57	36.3	497	5	O18192	O18192 caenorhabdi
44	57	36.3	500	5	Q27487	Q27487 caenorhabdi
45	57	36.3	512	5	Q9XV24	Q9xv24 caenorhabdi
46	57	36.3	524	5	Q9XV25	Q9xv25 caenorhabdi
47	57	36.3	702	16	Q98J51	Q98j51 rhizobium l
48	56	35.7	527	11	Q91X12	Q91x12 mus musculu
49	54	34.4	506	2	Q9EW74	Q9ew74 streptomyce
50	54	34.4	535	2	Q59141	Q59141 aeromonas s
51	54	34.4	567	16	Q9KRQ1	Q9krq1 vibrio chol
52	53	33.8	112	2	Q9EVN9	Q9evn9 listeria mo
53	52	33.1	112	5	Q9XVH1	Q9xvh1 caenorhabdi
54	52	33.1	488	16	Q926X0	Q926x0 listeria in
55	51	32.5	701	2	O60105	O60105 xanthomonas
56	51	32.5	702	2	Q9RB75	Q9rb75 xanthomonas
57	51	32.5	787	2	O69511	O69511 mycobacteri
58	51	32.5	819	10	Q9C5X8	Q9c5x8 arabidopsis
59	51	32.5	856	15	Q90CH3	Q90ch3 human immun
60	51	32.5	1030	2	Q9XD99	Q9xd99 clostridium
61	50	31.8	252	16	Q9KBV5	Q9kbv5 bacillus ha
62	50	31.8	327	16	Q92Y18	Q92y18 rhizobium m
63	50	31.8	854	15	Q97016	Q97016 human immun
64	50	31.8	854	15	Q72744	Q72744 human immun
65	50	31.8	1216	4	Q16827	Q16827 homo sapien
66	49	31.2	167	2	Q93079	Q93079 rhodococcus
67	49	31.2	492	10	Q9M502	Q9m502 helianthus
68	49	31.2	492	10	O24107	O24107 nicotiana g
69	49	31.2	823	5	Q22822	Q22822 caenorhabdi
70	48.5	30.9	1271	9	Q94MW9	Q94mw9 streptococc
71	48	30.6	318	12	Q9J9S7	Q9j9s7 human coxsa
72	48	30.6	318	12	Q9J9S6	Q9j9s6 human coxsa
73	48	30.6	318	12	Q9J9R1	Q9j9r1 human coxsa
74	48	30.6	318	12	Q9J9R0	Q9j9r0 human coxsa
75	48	30.6	318	12	Q9J9Q9	Q9j9q9 human coxsa
76	48	30.6	318	12	Q9J9Q8	Q9j9q8 human coxsa
77	48	30.6	351	4	Q96S51	Q96s51 homo sapien
78	48	30.6	474	2	Q59142	Q59142 aeromonas s
79	48	30.6	506	5	Q9VLI6	Q9vli6 drosophilla
80	48	30.6	682	10	Q9FX72	Q9fx72 arabidopsis
81	48	30.6	763	5	Q22873	Q22873 caenorhabdi
82	48	30.6	1213	10	Q9LHK2	Q9lhk2 arabidopsis
83	48	30.6	1226	11	Q9ERM5	Q9erm5 mus musculu
84	48	30.6	1297	3	Q9Y817	Q9y817 schizosacch
85	48	30.6	4601	5	Q9V383	Q9v383 drosophilla
86	47.5	30.3	150	2	O52168	O52168 salmonella
87	47.5	30.3	1428	2	Q9F922	Q9f922 serratia en
88	47	29.9	50	12	Q9J9V6	Q9j9v6 human coxsa
89	47	29.9	161	2	Q9L7Y4	Q9l7y4 pseudomonas





```
RP SEQUENCE FROM N.A.
RA Pua E.C.;
RT "Molecular cloning and expression of catalase genes in Brassica
   juncea.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC EMBL; AF104452; AAD17934.1; -.
CC HSSP; P21179; ICF9.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56828 MW; 3AD5D73AF1AD7C3C CRC64;

Query Match 47.8%; Score 75; DB 10; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.0057;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TTAFCGAPVWDDNNVITAGPRG 29
Db 18 TTNSGAPVWNNSSMTVGPRG 38

RESULT 5
Q9SYT7 PRELIMINARY; PRT; 492 AA.
AC Q9SYT7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13; Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CAT3.
OS Brassica juncea (Leaf mustard) (Indian mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3707;
RN [1]
RP SEQUENCE FROM N.A.
RA Pua E.C.;
RT "Molecular cloning and expression of catalase genes in Brassica
   juncea.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC EMBL; AF104453; AAD17935.1; -.
CC HSSP; P00432; 4BLC.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56915 MW; F17F193E2E988F5 CRC64;

Query Match 47.8%; Score 75; DB 10; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.0057;
Matches 13; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 9 TTAFCGAPVWDDNNVITAGPRG 29
Db 18 TTNSGAPVWNNSSMTVGPRG 38
```

```
RESULT 6
Q9M4X3 PRELIMINARY; PRT; 492 AA.
ID Q9M4X3;
AC Q9M4X3;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15; Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CAT.
OS Raphanus sativus (Radish).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Raphanus.
OX NCBI_TaxID=3726;
RN [1]
RP SEQUENCE FROM N.A.
RA Kwon S.-I.; An C.-S.;
RT "Characterization and expression of the catalase gene from small
   radish.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC EMBL; AF248491; AAF71742.1; -.
CC HSSP; P21179; ICF9.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56817 MW; 03E38F08D038276E CRC64;

Query Match 47.8%; Score 75; DB 10; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.0057;
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 9 TTAFCGAPVWDDNNVITAGPRG 29
Db 18 TTNSGAPVWNNSSMTVGPRG 38

RESULT 7
Q940G2 PRELIMINARY; PRT; 492 AA.
ID Q940G2;
AC Q940G2;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19; Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE CATALASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A.; Karlin-Neumann G.; Nguyen M.; Lam B.; Miranda M.;
RA Palm C.J.; Bowser L.; Jones T.; Banh J.; Carninci P.; Chen H.;
RA Cheuk R.; Chung M.K.; Hayashizaki Y.; Ishida J.; Kamiya A.; Kawai J.;
RA Kim C.; Lin J.; Liu S.X.; Narusaka M.; Pham P.K.; Sakano H.;
RA Sakurai T.; Satou M.; Seki M.; Shinn P.; Yamada K.; Shinozaki K.;
RA Ecker J.; Theologis A.; Davis R.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY054663; AAK96854.1; -.
SQ SEQUENCE 492 AA; 56888 MW; AB6225D56561FDABB CRC64;
```



```
AC Q9RG14;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN KATA.
OS Actinobacillus actinomycetemcomitans (Haemophilus
OS actinomycetemcomitans).
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Actinobacillus.
OX NCBI_TaxID=714;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y4NAL;
RX MEDLINE=20042351; PubMed=10572134;
RA Thompson V.J., Bhattacharjee M.K., Fine D.H., Derbyshire K.M.,
RA Figurski D.H.;
RT Direct Selection of IS903 Transposon Insertions by Use of a Broad-
RT Host-Range Vector: Isolation of Catalase-Deficient Mutants of
RT Actinobacillus actinomycetemcomitans.;
RL J. Bacteriol. 181:7298-7307(1999).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC EMBL; AF162654; AAF17882.1; -.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase_1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 484 AA; 54961 MW; D0E523AB2557D8CB CRC64;

Query Match 47.1%; Score 74; DB 2; Length 484;
Best Local Similarity 66.7%; Pred. No. 0.0079;
Matches 14; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 9 TTAAGAPVVDNNVITAGPRG 29
DB 15 TTAAGAPVVDNNTWSAGPRG 35

RESULT 12
Q9L4S1
AC Q9L4S1; PRELIMINARY; PRT; 505 AA.
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12600;
RA Sanz R.;
RL Thesis (1998), UNIV. COMPLUTENSE MADRID.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 12600;
RX MEDLINE=20170679; PubMed=10708385;
RA Sanz R., De la Fuente R.;
RT "Catalase deficiency in Staphylococcus aureus subsp. anaerobius is
RT associated with natural loss of function mutations within the
RT structural gene.";
RL Microbiology 146:465-475(2000).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
```

```
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
DR EMBL; AJ000472; CAB76839.1; -.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 505 AA; 58324 MW; BEEB05FBB0AE172F CRC64;

Query Match 47.1%; Score 74; DB 2; Length 505;
Best Local Similarity 55.2%; Pred. No. 0.0083;
Matches 16; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVNKDVKQTAFGAPVWDDNNVITAGPRG 29
DB 1 MSQDDKKLTGVFGHPVSDRENSMTAGPRG 29

RESULT 13
Q99UE2
ID Q99UE2 PRELIMINARY; PRT; 507 AA.
AC Q99UE2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN KATA OR SALL70 OR SAV1334.
OS Staphylococcus aureus (Strain N315), and
OS Staphylococcus aureus (Strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879; 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
DR EMBL; AP003133; BAB42427.1; -.
DR EMBL; AP003362; BAB57496.1; -.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Complete proteome; Heme; Hydrogen peroxide; Iron; Oxidoreductase;
KW Peroxidase.
SQ SEQUENCE 507 AA; 58612 MW; F6B9EE770E2DCBCC CRC64;

Query Match 47.1%; Score 74; DB 16; Length 507;
Best Local Similarity 55.2%; Pred. No. 0.0083;
Matches 16; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
```

```
OY 1 MYNKDVKOTAFGAPVWDNNVITAGPRG 29
   | : | | | | | | | | | | | | | |
Db 3 MSQDKKLTGVGHPVSDRENSMTAGPRG 31

RESULT 14
Q9SDS9
ID Q9SDS9 PRELIMINARY; PRT; 492 AA.
AC Q9SDS9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CAT1.
OS Zantedeschia aethiopica (White calla lily).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Araceae; Zantedeschia.
OX NCBI_TaxID=69721;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Lino-Neto T., Tavares R.M., Palme K., Pais M.S.S.:
RT "Expression of catalase during senescence and regreening of
RT Zantedeschia aethiopica spathe.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; AF207906; AAF19965.1; -.
DR HSSP; F21179; 1QF7.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56968 MW; DC9F3EB1392DE1CF CRC64;

Query Match 46.5%; Score 73; DB 10; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.011;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 9 TTAFGAPVWDNNVITAGPRG 29
   | | | | | | | | | | | | | |
Db 18 TTNSGAPVWNNNSITVGARG 38

RESULT 15
O09484
ID O09484 PRELIMINARY; PRT; 492 AA.
AC O09484;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC Thome H., Happe T.;
RA Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; Y13220; CAA73663.1; -.
DR HSSP; P00432; 4BLC.
```

```
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56227 MW; 3711B90126911C32 CRC64;

Query Match 45.9%; Score 72; DB 10; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.016;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 9 TTAFGAPVWDNNVITAGPRG 29
   | | | | | | | | | | | | | |
Db 18 TTNSGAPVWNNNSMTVGTGRG 38

RESULT 16
O22472
ID O22472 PRELIMINARY; PRT; 493 AA.
AC O22472;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CAT.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadaceae; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CC-125 WILD TYPE MT+ 137C;
RA Kim J.Y., Park K.Y., Lee K.O., Lee S.H.;
RT "Isolation and characterization of a cDNA encoding a catalase from
RT Chlamydomonas reinhardtii (Accession No. AF016902) (PCR 00-016).";
RL Plant Physiol. 122:293-293(2000).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; AF016902; AAB70006.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 493 AA; 55943 MW; 2B74B2BBA2CEE773 CRC64;

Query Match 45.9%; Score 72; DB 10; Length 493;
Best Local Similarity 61.9%; Pred. No. 0.016;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 9 TTAFGAPVWDNNVITAGPRG 29
   | | | | | | | | | | | | | |
Db 18 TTNSGAPVWNNNSMTVGTGRG 38

RESULT 17
O9AQO9
ID O9AQO9 PRELIMINARY; PRT; 483 AA.
AC O9AQO9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN KATA.
```

Query Match 45.2%; Score 71; DB 2; Length 484;  
Best Local Similarity 65.2%; Pred. No. 0.023;  
Matches 15; Conservative 2; Mismatches 6; Indels

```

RESULT 20
O22640
ID O22640 PRELIMINARY; PRT; 491 AA.
AC
CC O22640;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
DE CAT1.
GN Raphanus sativus (Radish).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Raphanus.
OX NCBI_TaxID=3726;
[1]
RN
RP SEQUENCE FROM N.A.
RA
RL I1 K-S, An C-S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RC
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
DR EMBL; AF031318; AAB86582.2; -.
DR HSSP; P21179; 1QF7.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase.1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE.1; 1.

```

DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 491 AA; 56716 MW; 7C20151C3BB16664 CRC64;

Query Match 44.6%; Score 70; DB 10; Length 491;  
 Best Local Similarity 60.0%; Pred. No. 0.033;  
 Matches 12; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 10 TAFGAPVWDDNNVITAGPRG 29  
 | |||||:::|:| |||||  
 Db 18 TNSGAPVWNNSSMTVGPRG 37

RESULT 21  
 Q9SB22 PRELIMINARY; PRT; 492 AA.

AC Q9SB22;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE (EC 1.11.1.6).  
 GN CATB.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;  
 RT "Evolutionary relationship of plant catalase genes inferred from exon-intron structures: isozyme divergence after the separation of monocots and dicots.";  
 RL Theor. Appl. Genet. 97:9-19(1998).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 DR EMBL; D64013; BAA34204.1; -;  
 DR HSSP; P21179; IQF7.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 492 AA; 56575 MW; 8C49D4FE45993FC8 CRC64;

Query Match 44.6%; Score 70; DB 10; Length 492;  
 Best Local Similarity 57.1%; Pred. No. 0.033;  
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAGFAPVWDDNNVITAGPRG 29  
 || |||||:::|:| |||||  
 Db 18 TTNSGAPVWNNSSALTIVGERG 38

RESULT 22  
 Q9RJK9 PRELIMINARY; PRT; 487 AA.

AC Q9RJK9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE (EC 1.11.1.6).  
 GN KATA.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomyces.

OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Murphy L., Harris D.;  
 RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=97000351; PubMed=8843436;  
 RX Redendach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 DR EMBL; AL121855; CAB58320.1; -;  
 DR HSSP; P42321; 2CAE.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 487 AA; 55116 MW; 9D3334889EAF60B7 CRC64;

Query Match 43.9%; Score 69; DB 2; Length 487;  
 Best Local Similarity 51.9%; Pred. No. 0.046;  
 Matches 14; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 3 NKDVKQTTAFGAPVWDDNNVITAGPRG 29  
 | || |||||:::|:| |||||  
 Db 4 NNOKPLTTVAGAPVDPNQNLSLTSGPRG 30

RESULT 23  
 Q9ZWH7 PRELIMINARY; PRT; 492 AA.  
 AC Q9ZWH7;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE.  
 GN OSCATC.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kim S., Park S., Lee Y.;  
 RT "Cloning and expression of a new rice catalase gene.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB020502; BAA34714.1; -;  
 DR HSSP; P00432; 4BLC.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 2.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 SQ SEQUENCE 492 AA; 56175 MW; A8D3D10D309443F7 CRC64;

```

Query Match      43.9%; Score 69; DB 10; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.046;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29
   :| | | | | :| | | | |
Db 18 STNSGAPVWNNNSLTGSRG 38

RESULT 24
ID Q9ZRI9 PRELIMINARY; PRT; 492 AA.
AC Q9ZRI9;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATC.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriactoidae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LEAF;
RA Iwamoto M., Maekawa M., Saito A., Higo H., Higo K.;
RT "Evolutionary relationship of plant catalase genes inferred from exon-
RT intron structures: isozyme divergence after the separation of monocots
RT and dicots.";
RL Theor. Appl. Genet. 97:9-19(1998).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC EMBL; D86611; BAA34205.1; -
DR HSSP; P00432; 4BLC.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase: 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase: 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56806 MW; B14CE64C275213C0 CRC64;

Query Match      43.9%; Score 69; DB 10; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.046;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29
   :| | | | | :| | | | |
Db 18 STNSGAPVWNNNSLTGSRG 38

RESULT 25
ID Q9FEK5 PRELIMINARY; PRT; 492 AA.
AC Q9FEK5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CAT.
OS Hevea brasiliensis (Para rubber tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaeae; Hevea.
OX NCBI_TaxID=3981;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Kongsawadworakul P., Pujade-Renaud V., Narangajavana J., Montoro P.,
RA Lacroite R., Chrestin H.;
RT "Cloning and characterization of a catalase cDNA from the latex cells
RT of Hevea brasiliensis. Short and long term effect of ethylene on the
RT corresponding catalase and the MnSOD genes expression in the latex.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF151368; AAG43363.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase: 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase: 1.
DR PROSITE; PS00437; CATALASE_1; 1.
KW Oxidoreductase; Peroxidase.
SQ SEQUENCE 492 AA; 56705 MW; 73C4361F837EEB8C CRC64;

Query Match      43.3%; Score 68; DB 10; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.065;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29
   || | | | | :| | | | |
Db 18 TTNSGAPVWNNNSLTGARG 38

RESULT 26
Q9LDS9 PRELIMINARY; PRT; 1013 AA.
ID Q9LDS9;
AC Q9LDS9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F2D10 from chromosome
RT I.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F5M15 from chromosome
RT I.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.

```

RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altati H., Bei B., Chin C., Chio J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.,  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
CC PEROXIDE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) - O(2) + 2 H(2)O.  
DR EMBL; AC069251; AAF80611.1; -;  
DR EMBL; AC027665; AAF79625.1; -;  
DR HSSP; P00432; 4BLC.  
DR InterPro; IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 2.  
DR PRINTS; PR00067; CATALASE.  
DR PRODOM; PD000510; Catalase; 2.  
DR PROSITE; PS00437; CATALASE\_1; 2.  
DR PROSITE; PS00438; CATALASE\_2; 1.  
DR Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
SQ SEQUENCE 1013 AA; 116681 MW; 604797360DB8486D CRC64;

Query Match 43.3%; Score 68; DB 10; Length 1013;  
Best Local Similarity 57.1%; Pred. No. 0.15;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
||| |||||:::|:| |||  
Db 539 TTNSGAPVWNNSSLTGTRG 559

RESULT 27  
Q9SW99 ID Q9SW99 PRELIMINARY; PRT; 492 AA.  
AC Q9SW99;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CATALASE CATI (EC 1.11.1.6).  
OS Manihot esculenta (Cassava) (Manioc).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Malpighiales; Euphorbiaceae; Manihot.  
OX NCBI\_TaxID=3983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-STORAGE ROOT;  
RX MEDLINE-21210978; PubMed-11311947;  
RA Reilly K., Han Y., Tohme J., Beeching J.R.;  
RT "Isolation and characterisation of a cassava catalase expressed during  
RT post-harvest physiological deterioration";  
RL Biochim. Biophys. Acta 1518:317-323(2001).  
DR EMBL; AF170272; AAD50974.1; -;  
DR HSSP; P21179; 1CF9.  
DR InterPro; IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR PRODOM; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; 1.  
KW Oxidoreductase; Peroxidase.  
SQ SEQUENCE 492 AA; 57173 MW; 11C75B456CB6C6F8 CRC64;

Query Match 42.7%; Score 67; DB 10; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.093;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
||| |||||:::|:| |||  
Db 18 TTDAGAPVWNNSSMTGTRG 38

RESULT 28  
Q94EV9 ID Q94EV9 PRELIMINARY; PRT; 492 AA.  
AC Q94EV9;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CATALASE.  
GN CAT.

OS Suaeda maritima subsp. salsa.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Suaeda.  
OX NCBI\_TaxID=126914;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wang P.P., Ma C.L., Sun Y.F., Zhao Y.X., Zhang H.;  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF390210; AAK67359.1; -;  
SQ SEQUENCE 492 AA; 56788 MW; 8532135DE66DA3B6 CRC64;

Query Match 42.7%; Score 67; DB 10; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.093;  
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
||| |||||:::|:| |||  
Db 18 TANSAGPVWNNSSLTGVRG 38

RESULT 29  
Q9FAZ3 ID Q9FAZ3 PRELIMINARY; PRT; 509 AA.  
AC Q9FAZ3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CATALASE VKTA.  
GN VKTA.  
OS Vibrio rumoiensis.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=76258;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S-1;  
RA Ichise N., Morita N., Hoshino T., Kawasaki K., Yumoto I., Okuyama H.;  
RT "Cloning of the Complete Catalase Gene from the Hydrogenperoxide-  
RT resistant Bacterium Vibrio rumoiensis S-1 and Its Subcellular  
RT localization";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB030821; BAB12412.1; -;  
DR HSSP; P42321; 2CAE.  
DR InterPro; IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR PRODOM; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; UNKNOWN\_1.  
SQ SEQUENCE 509 AA; 57659 MW; CE3A2B4FD1C7DEE8 CRC64;

Query Match 42.0%; Score 66; DB 2; Length 509;  
Best Local Similarity 61.9%; Pred. No. 0.14;  
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
||| |||||:::|:| |||  
Db 14 TTDGAPVWNNSSLTGPRG 34

RESULT 30





RA Whitehead S., Spratt B.G., Barrell B.G.:  
 RT \*Complete DNA sequence of a serogroup A strain of *Neisseria*  
 RT meningitidis Z2491.";  
 RL Nature 404:502-506(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC58 / SEROGROUP B;  
 RX MEDLINE=20175755; PubMed=10710307;  
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
 RA Mason T., Cleckos A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,  
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
 RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,  
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.:  
 RT \*Complete genome sequence of *Neisseria meningitidis* serogroup B strain  
 RT MC58.";  
 RL Science 287:1809-1815(2000).  
 DR EMBL; AL162752; CAB83367.1; -;  
 DR EMBL; AE002379; AAF40672.1; -;  
 DR HSSP; P42321; 2CAE.  
 DR TIGR; NME0216; -;  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; UNKNOWN\_1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 DR Oxidoreductase; Peroxidase; Complete proteome.  
 KW SEQUENCE 504 AA; 57167 MW; 26858B47446F3775 CRC64;  
 SQ

Query Match 40.8%; Score 64; DB 16; Length 504;  
 Best Local Similarity 70.8%; Pred. No. 0.27;  
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 13 GAPVDDNNVITAGPRG 29  
 ||| | : : |||||  
 Db 16 GAPVADNNSLTAGPRG 32

RESULT 34  
 Q918V5 ID Q918V5 PRELIMINARY; PRT; 526 AA.  
 AC Q918V5  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE (EC 1.11.1.6).  
 GN CAT.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gerhard G.S., Kaufman E.J., Grundy M.A.:  
 RT \*Molecular cloning and sequence analysis of the *Danio rerio* catalase  
 RT gene.";  
 RL Comp. Biochem. Physiol. 127:447-457(2000).  
 CC -I- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE (BY SIMILARITY).  
 CC -I- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 DR EMBL; AF170069; AAF89686.1; -;  
 DR HSSP; P00432; 4BLC.  
 DR ZFIN; ZDB-GENE-000210-20; cat.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.

DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 526 AA; 59542 MW; 9FA05D42737662A4 CRC64;  
 RP

Query Match 40.8%; Score 64; DB 13; Length 526;  
 Best Local Similarity 57.1%; Pred. No. 0.29;  
 Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
 || | | : : |||||  
 Db 28 TTGAGVPIGDKLNAMTAGPRG 48

RESULT 35  
 P77924 ID P77924 PRELIMINARY; PRT; 506 AA.  
 AC P77924  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE ISOZYME.  
 GN KATA.  
 OS Pseudomonas fluorescens.  
 OG Plasmid pAM10.6.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CB36;  
 RX MEDLINE=21318967; PubMed=11425481;  
 RA Peters M., Heinaru A., Nark A.;  
 RT \*Plasmid-encoded catalase KATA, the main catalase of *Pseudomonas*  
 RT fluorescens strain CB36.";  
 RL FEMS Microbiol. Lett. 200:235-240(2001).  
 DR EMBL; U72068; AAB17009.1; -;  
 DR HSSP; P42321; 2CAE.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Plasmid.  
 SQ SEQUENCE 506 AA; 57324 MW; F1EAA728C5D41CBE CRC64;  
 RP

Query Match 39.5%; Score 62; DB 2; Length 506;  
 Best Local Similarity 61.9%; Pred. No. 0.55;  
 Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
 || |||| | : : |||||  
 Db 14 TTGAPVVDNQNMTAGARG 34

RESULT 36  
 Q9CPK5 ID Q9CPK5 PRELIMINARY; PRT; 484 AA.  
 AC Q9CPK5  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE HKTE.  
 GN HKTE OR PM0032.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 OX NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM70;

RX MEDLINE-21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of *Pasteurella multocida* pm70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 DR EMBL; AE006036; AA02116.1; -.  
 DR HSSP; P42321; 2CAE.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR PRODOM; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 484 AA; 55127 MW; 5BE648D3610A6420 CRC64;

Query Match 38.98; Score 61; DB 16; Length 484;  
 Best Local Similarity 47.6%; Pred. No. 0.74;  
 Matches 10; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDDNNVITAGPRG 29  
 I I I I I : : : : : I I I I I  
 Db 14 TNAAGAPIVENDNTWSAGPKG 34

## RESULT 37

ID Q9C1Q7 PRELIMINARY; PRT; 349 AA.  
 AC Q9C1Q7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
 DE UNKNOWN PROTEIN.  
 GN YCJ1.  
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Lactococcus.  
 OX NCBI\_TaxID=1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-IL1403;  
 RX MEDLINE-21235186; PubMed=11337471;  
 RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,  
 RT Weissenbach J., Ehrlich S.D., Sorokin A.;  
 RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*  
 RL *lactis* ssp. *lactis* IL1403.";  
 RL Genome Res. 11:731-753(2001).  
 DR EMBL; AE006267; AA04397.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 349 AA; 37380 MW; 4A71A677ECE3D3A4 CRC64;

Query Match 38.2%; Score 60; DB 16; Length 349;  
 Best Local Similarity 56.5%; Pred. No. 0.73;  
 Matches 13; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 7 KQTTAFGA--PVWDDNNVITAGP 27  
 I I I I I : I I I I I I  
 Db 256 KWTTFGAKESIWDDNTVIGTTP 278

## RESULT 38

ID Q9M504 PRELIMINARY; PRT; 492 AA.  
 AC Q9M504;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE (EC 1.11.1.6).  
 GN CATA2  
 OS Helianthus annuus (Common sunflower).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;

OC Heliantheae; Helianthus.  
 OX NCBI\_TaxID=4232;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. SPANNERS ALL2WECK; TISSUE-COTYLEDONS;  
 RA Heinze M., Baur B., Eising R.;  
 RT "Biogenesis of catalase forming the crystalline inclusions in  
 RL peroxisomes of sunflower cotyledons.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE (BY SIMILARITY).  
 CC CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 DR EMBL; AF243517; AAF61732.1; -.  
 DR HSSP; P00432; 4BLC.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR PRODOM; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 492 AA; 56931 MW; AADF2B516C130163 CRC64;

Query Match 37.6%; Score 59; DB 10; Length 492;  
 Best Local Similarity 47.6%; Pred. No. 1.5;  
 Matches 10; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDDNNVITAGPRG 29  
 I I I I I : : : : : I I I I I  
 Db 18 TTYGAPVYNDSSLTGVTRG 38

## RESULT 39

ID O81336 PRELIMINARY; PRT; 492 AA.  
 AC O81336;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE CATALASE (EC 1.11.1.6).  
 OS Mesembryanthemum crystallinum (Common ice plant).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Caryophyllaceae; Aizoaceae; Mesembryanthemum.  
 OX NCBI\_TaxID=3544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RA Michalowski C.B., Quigley-Landreau F., Bohnert H.J.;  
 RT "Different isoforms of catalase are expressed in leaves and roots of  
 RL the common ice plant.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE (BY SIMILARITY).  
 CC CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 DR EMBL; AF069319; AAC19397.1; -.  
 DR HSSP; P21179; 1CF9.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR PRODOM; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.  
 SQ SEQUENCE 492 AA; 56886 MW; 6DD5DB0AFAAD5C60 CRC64;

Query Match 37.6%; Score 59; DB 10; Length 492;  
 Best Local Similarity 52.4%; Pred. No. 1.5;  
 Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```

QY 9 TTAFGAPVWDNNVNITAGPRG 29
   |||:::|:|:|
Db 18 TTNSGAPVYNNSSSLVTGTRG 38

RESULT 40
P77948
ID P77948 PRELIMINARY; PRT; 488 AA.
AC P77948;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE CATALASE (EC 1.11.1.6).
GN CATA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC10147;
RX MEDLINE=97334128; PubMed=9190825;
RA Cho Y.H., Roe J.H.;
RT "Isolation and expression of the cata gene encoding the major
RT vegetative catalase in Streptomyces coelicolor muller.";
RL J. Bacteriol. 179:4049-4052(1997).
CC -I- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE (BY SIMILARITY).
CC -I- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
DR EMBL; X96981; CAA65681.1; -.
DR HSSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Heme; Hydrogen peroxide; Iron; Oxidoreductase; Peroxidase.
SQ SEQUENCE 488 AA; 5540 MW; 323633B73E46C7C8 CRC64;

Query Match 36.9%; Score 58; DB 2; Length 488;
Best Local Similarity 46.2%; Pred. No. 2.1;
Matches 12; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 3 NKDVKQTTAFGAPVWDNNVNITAGPR 28
   |||:::|:|:|
Db 4 NNQKPLTTVAGPPVDPNQNLSLTSGR 29

```

Search completed: August 22, 2002, 07:45:26  
Job time: 145 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:43:01 ; Search time 41.42 Seconds

(without alignments)  
91.885 Million cell updates/sec

Title: US-09-488-737-1  
Perfect score: 123  
Sequence: 1 MKEKFNRTKPHVNICTIGHVDH 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database :

SPTREMBL19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	112	91.1	397	16	Q97EH5
2	111	90.2	394	16	Q9KV37
3	111	90.2	394	16	Q9KUZ6
4	109	88.6	395	16	Q92716
5	107	87.0	394	16	Q9K117
6	107	87.0	394	16	Q9JRI5
7	107	87.0	395	2	Q9R420
8	107	87.0	395	2	Q9RHI2
9	107	87.0	395	2	Q9RHI1
10	107	87.0	395	2	Q9RHI0
11	107	87.0	395	2	Q9RHH9
12	107	87.0	396	16	Q99QM0
13	106	86.2	394	2	Q933R9
14	106	86.2	394	16	Q92GM4
15	106	86.2	397	2	Q9RG55
16	106	86.2	397	16	Q9HWD3

17	106	86.2	398	16	Q97PV3
18	105	85.4	391	2	Q9F9S8
19	105	85.4	397	2	Q9RG53
20	105	85.4	397	2	Q93T39
21	105	85.4	401	2	Q9XD38
22	102	82.9	358	2	Q9X6G1
23	102	82.9	391	16	Q925Y6
24	102	82.9	467	10	Q9SEF8
25	100	81.3	393	2	Q9RAJ9
26	99	80.5	395	16	Q9CEI0
27	98	79.7	410	8	Q9TLV8
28	97	78.9	394	16	Q99W61
29	97	78.9	409	8	Q9TJQ8
30	97	78.9	437	3	Q9C2P1
31	96	78.0	457	10	Q0450
32	96	78.0	474	10	Q9AXU2
33	96	78.0	476	10	Q93Y56
34	96	78.0	485	10	Q43364
35	95	77.2	300	5	Q95TV3
36	95	77.2	489	5	Q9V607
37	94	76.4	119	8	Q78381
38	94	76.4	396	16	Q9P9Q9
39	94	76.4	398	2	Q9AIG6
40	94	76.4	401	5	Q26114
41	94	76.4	401	8	Q9TMN9
42	94	76.4	401	8	Q78325
43	91	74.0	454	10	Q9ZT91
44	91	74.0	471	10	Q39206
45	90	73.2	452	10	Q9F0Z6
46	89	72.4	496	5	Q19072
47	87	70.7	396	16	Q98QG1
48	86	69.9	403	8	Q33451
49	78	63.4	410	5	Q25820
50	72	58.5	456	5	Q9V4M7
51	66	53.7	395	2	Q9S513
52	66	53.7	427	17	Q979T1
53	65	52.8	50	1	Q93633
54	65	52.8	88	1	Q93638
55	65	52.8	411	17	Q978W8
56	65	52.8	415	17	Q980A5
57	64	52.0	422	17	Q975N8
58	62	50.4	414	17	Q9HNK9
59	62	50.4	439	5	Q76135
60	62	50.4	453	5	P91150
61	61	49.6	435	17	Q976B1
62	60	48.8	468	5	Q25099
63	59	48.0	565	17	Q29514
64	58	47.2	742	3	Q9HG15
65	57	46.3	85	2	Q53774
66	57	46.3	435	10	Q82788
67	55	44.7	13	2	Q47693
68	54.5	44.3	63	10	Q42012
69	54.5	44.3	68	10	P94009
70	54.5	44.3	81	10	P94008
71	54.5	44.3	121	10	P94007
72	54.5	44.3	447	10	Q9ZRP9
73	54.5	44.3	447	10	P93769
74	54.5	44.3	449	10	Q9C514
75	54.5	44.3	449	10	Q39093
76	54.5	44.3	449	10	P9AS09
77	54.5	44.3	449	10	Q94AD0
78	54.5	44.3	465	5	Q9U600
79	54.5	44.3	967	10	Q9LNL3
80	54	43.9	97	10	Q42336
81	54	43.9	471	10	Q49736
82	53	43.1	389	16	Q9K7L3
83	53	43.1	540	17	Q9HPE4
84	53	43.1	573	5	Q45622
85	52	42.3	16	6	Q9TRQ9
86	52	42.3	131	5	Q9N9D5
87	52	42.3	132	5	Q9NA07
88	52	42.3	465	10	O64490
89	52	42.3	465	10	Q9ZV47

Q97PV3	streptococ
Q9F9S8	edta-degrad
Q9RG53	streptococ
Q93T39	streptomyc
Q9XD38	leptospi
Q9X6G1	coxiella bu
Q925Y6	rhizobium m
Q9SEF8	oryza sativ
Q9RAJ9	moraxella s
Q9CEI0	lactococcus
Q9TLV8	cyanidium c
Q99W61	staphylococ
Q9TJQ8	prototheca
Q9C2P1	neurospora
Q0450	nicotiana s
Q9AXU2	pelargonium
Q93Y56	arabidopsis
Q43364	nicotiana s
Q95TV3	drosophila
Q9V607	drosophila
Q78381	theileria a
Q9P9Q9	xylella fas
Q9AIG6	candidate
Q26114	toxoplasma
Q9TMN9	toxoplasma
Q78325	toxoplasma
Q9ZT91	arabidopsis
Q39206	arabidopsis
Q9F0Z6	zea mays (m
Q19072	caenorhabdi
Q98QG1	mycoplasma
Q33451	eimeria ten
Q25820	plasmodium
Q9V4M7	drosophila
Q9S513	loofah witec
Q979T1	thermoplasma
Q93633	methanococ
Q93638	methanococ
Q978W8	thermoplasma
Q980A5	sulfolobus
Q975N8	sulfolobus
Q9HNK9	halobacteri
Q76135	caenorhabdi
P91150	caenorhabdi
Q976B1	sulfolobus
Q25099	hydra magni
Q29514	archaeoglob
Q9HG15	yarrowia li
Q53774	stigmatella
Q82788	blastocysti
Q47693	escherichia
Q42012	arabidopsis
P94009	arabidopsis
P94008	arabidopsis
P94007	arabidopsis
Q9ZRP9	malus domes
Q93769	nicotiana t
Q9C514	arabidopsis
Q39093	arabidopsis
P9AS09	arabidopsis
Q94AD0	arabidopsis
Q9U600	anisakis si
Q9LNL3	arabidopsis
Q42336	arabidopsis
Q49736	arabidopsis
Q9K7L3	bacillus ha
Q9HPE4	halobacteri
Q45622	caenorhabdi
Q9TRQ9	sus scrofa
Q9N9D5	scoliotteri
Q9NA07	drosophila
O64490	arabidopsis
Q9ZV47	arabidopsis

NCBI\_TaxID=666;  
[1]  
SEQUENCE FROM N.A.  
RN STRAIN=EL TOR N16961 / SEROTYPE O1;  
RP  
RC MEDLINE=20406833; PubMed=10952301;  
RX  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;

0  
4  
4  
2  
4  
. .  
:  
)  
7  
)  
:  
)  
:  
2  
  
3  
4  
)  
4  
0  
:  
)  
  
)  
:  
)  
  
4

CC DR EMBL; A5004120; AAF93494.1; -.  
DR HSSP; P02990; 1ETU.  
DR TIGR; VC0321; -.  
DR InterPro: IPR000795; GTP\_EFTU.  
DR PRINTS; PR00315; ELONGATNFCT.  
DR PROSITE; PS00301; EFACOR\_GTP; 1.  
DR Complete proteome: GTP-binding.  
SQ SEQUENCE 394 AA; 43185 MW; A047E98A84B9A40 CRC64;

RA Dodson R.J., Hail D.H., Hickey E.N., Peterson J.D., Umayam L.A.,  
RA Gill S.P., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bagg S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*

of the cholera pathogen *Vibrio*

```
CC      - 1- SIMILARITY BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR      EMBL; AE004124; AAF93535.1; -.
DR      HSSP; P02990; 1ETU.
DR      TIGR; VC0362; -.
DR      InterPro; IPR000795; GTP_EFTU.
DR      PRINTS; PR00315; ELONGATNFCT.
DR      PROSITE; PS00301; EFACOR_GTP; 1.
DR      Complete proteome; gtp-binding.
SQ      SEQUENCE 394 AA; 43136 MW; BF0DE31A00618B36 CRC64;
```

Query Match 90.28; Score 111; DB 16; Length 394;

Best Local Similarity 90.5%; Pred. No. 8.5e-10;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNICTIGHVDH 22  
Db 3 KEKFNRTKPHVNICTIGHVDH 23

RESULT 4  
ID Q927I6 PRELIMINARY; PRT; 395 AA.  
AC Q927I6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TUPA PROTEIN  
GN TUPA OR LIN2802.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Listeria.  
OX NCBI\_TaxID=1642;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CLIP 11262 / SEROVAR 6A;  
RX PubMed=11679669;  
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,  
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,  
RA Gautier L., Goebel W., Gomez-Iopez N., Hain T., Hauf J., Jackson D.,  
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,  
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
RA Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,  
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
RA Varquez-Boland J.-A., Voss H., Wehland J., Cossart P.;  
RT "Comparative genomics of Listeria species.";  
RL Science 294:849-852(2001).  
DR EMBL; AL596173; CAC98028.1; -.  
DR Listlist; LIN2802; -.  
KW Complete proteome.  
SQ SEQUENCE 395 AA; 43353 MW; CE0FA4906D644C24 CRC64;

Query Match 88.6%; Score 109; DB 16; Length 395;  
Best Local Similarity 90.5%; Pred. No. 1.8e-09;  
Matches 19; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNICTIGHVDH 22  
Db 3 KEKFNRTKPHVNICTIGHVDH 23

RESULT 5  
ID Q9K1I7 PRELIMINARY; PRT; 394 AA.  
AC Q9K1I7;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE TRANSLATION ELONGATION FACTOR TU.  
GN NMB0139.  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
GN NMB0139.  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,

Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,  
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
RA Gill J., Scarlato V., Masignani V., Pizze M., Grandi G., Sun L.,  
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain  
MC58.";  
RL Science 287:1809-1815(2000).  
CC -|- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
DR EMBL; AE002372; AAF40598.1; -.  
DR HSSP; P02990; 1EFTU.  
DR TIGR; NMB0139; -.  
DR InterPro; IPR000795; GTP\_EFTU.  
DR PRINTS; PR00315; ELONGATNFT.  
DR PROSITE; PS00301; EFACFOR\_GTP; 1.  
KW Complete proteome; Elongation factor; GTP-binding.  
SQ SEQUENCE 394 AA; 42925 MW; 55762F03EC048563 CRC64;

Query Match 87.0%; Score 107; DB 16; Length 394;  
Best Local Similarity 85.7%; Pred. No. 3.8e-09;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNICTIGHVDH 22  
Db 3 KEKFNRTKPHVNICTIGHVDH 23

RESULT 6  
ID Q9JRI5 PRELIMINARY; PRT; 394 AA.  
AC Q9JRI5;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE ELONGATION FACTOR TU (TRANSLATION ELONGATION FACTOR TU).  
GN (TUFAL OR NMA0134 OR NMB0124) AND (TUF2 OR NMA0149).  
OS Neisseria meningitidis (serogroup A), and  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699; 491;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;  
RX MEDLINE=2022556; PubMed=10761919;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,  
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,  
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,  
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,  
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,  
RA Whitehead S., Spratt B.G., Barrell B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria  
meningitidis Z2491.";  
RL Nature 404:502-506(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,  
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,  
RA Gill J., Scarlato V., Masignani V., Pizze M., Grandi G., Sun L.,  
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain  
MC58.";  
RL Science 287:1809-1815(2000).  
CC -|- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
DR EMBL; AL162752; CAB83449.1; -.  
DR EMBL; AE002371; AAF40583.1; -.  
DR EMBL; AL162752; CAB83464.1; -.  
DR HSSP; P02990; 1EFTU.





```

ID Q9RH10 PRELIMINARY; PRT; 395 AA.
AC Q9RH10;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EF-TU.
GN EF-TU.
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Porphyromonadaceae;
OC Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33277;
RA Nagai A., Itoh N.;
RT "EF-Tu sequence of Porphyromonas gingivalis.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035465; BAA88139.1; -.
DR HSSP; P02990; 1EFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACFOR_GTP; 1.
KW GTP-binding.
SQ SEQUENCE 395 AA; 43669 MW; 1F2F39D3173ADA3F CRC64;

Query Match 87.0%; Score 107; DB 2; Length 395;
Best Local Similarity 85.7%; Pred. No. 3.8e-09;
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KEHFNRSKPHVNVGTIGHVDH 23

RESULT 11
Q9RH9 PRELIMINARY; PRT; 395 AA.
ID Q9RH9;
AC Q9RH9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE EF-TU.
GN EF-TU.
OS Bacteroides forsythus.
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
OC Bacteroides.
OX NCBI_TaxID=28112;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 43037;
RA Nagai A., Itoh N.;
RT "EF-Tu sequence of Bacteroides forsythus.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB035466; BAA88139.1; -.
DR HSSP; P02990; 1EFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACFOR_GTP; 1.
KW GTP-binding.
SQ SEQUENCE 395 AA; 43695 MW; 440B33AEBF39CEF CRC64;

Query Match 87.0%; Score 107; DB 2; Length 395;
Best Local Similarity 85.7%; Pred. No. 3.8e-09;
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KEHFNRSKPHVNVGTIGHVDH 23

RESULT 12
Q9QM0 PRELIMINARY; PRT; 396 AA.
ID Q9QM0;
AC Q9QM0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE TRANSLATION ELONGATION FACTOR EF-TU.
GN CC3199 AND CC1240.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kotonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ullerback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC - SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL; AE005984; AAK25161.1; -.
DR EMBL; AE005800; AAK23221.1; -.
DR HSSP; P02990; 1ETU.
DR TIGR; CC1240; -.
DR TIGR; CC3199; -.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACFOR_GTP; 1.
KW Complete proteome; Elongation factor; GTP-binding.
SQ SEQUENCE 396 AA; 43339 MW; 1F58829078B624D2 CRC64;

Query Match 87.0%; Score 107; DB 16; Length 396;
Best Local Similarity 90.5%; Pred. No. 3.8e-09;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KEKFERTKPHCNIGTIGHVDH 23

RESULT 13
Q93R9 PRELIMINARY; PRT; 394 AA.
ID Q93R9;
AC Q93R9;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN TUF.
OS Ureaplasma urealyticum (Ureaplasma urealyticum biotype 2).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;

```



Query Match 86.2%; Score 106; DB 16; Length 397;  
Best Local Similarity 85.7%; Pred. No. 5.6e-09;  
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTIGTIGHVDH 22  
||||| | |||||:|||||||

Db 3 KEKPERKPHVNTIGTIGHVDH 23

RESULT 17

Q97PV3 ID Q97PV3 PRELIMINARY; PRT; 398 AA.  
AC Q97PV3;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TRANSLATION ELONGATION FACTOR TU.  
GN SP1489.  
OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX STRAIN=TIGR4;  
RX MEDLINE=21357209; PubMed=11463916;  
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
RA Unayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
RT "Complete genome sequence of a virulent isolate of Streptococcus  
RT pneumoniae".  
RL Science 293:498-506(2001).  
DR EMBL; AE007444; AAK75581.1; -.  
DR TIGR; SP1489; -.  
DR InterPro; IPR000795; GTP\_EFTU.  
DR InterPro; IPR004161; GTP\_EFTU\_D2.  
DR InterPro; IPR004160; GTP\_EFTU\_D3.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
DR PRINTS; PR00315; ELONGATNFCT.  
DR PROSITE; PS00301; EFATOR.GTP; 1.  
KW Elongation factor; Complete proteome.  
SQ SEQUENCE 398 AA; 43970 MW; 5CB2F8F5FA5101A0 CRC64;

Query Match 86.2%; Score 106; DB 16; Length 398;  
Best Local Similarity 85.7%; Pred. No. 5.6e-09;  
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTIGTIGHVDH 22  
||||| | |||||:|||||||

Db 3 KEKDRSKPHVNTIGTIGHVDH 23

RESULT 18

Q9F9S8 ID Q9F9S8 PRELIMINARY; PRT; 391 AA.  
AC Q9F9S8;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE EFTU.  
GN EFTU.  
OS EDTA-degrading bacterium BNCL.  
OC Bacteria; Proteobacteria; alpha subdivision.  
OX NCBI\_TaxID=85561;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=BNCL.  
RX MEDLINE=21091958; PubMed=11157232;  
RA Bohuslavsek J., Payne J.W., Liu Y., Bolton H. Jr., Xun L.;  
RT "Cloning, Sequencing, and Characterization of a Gene Cluster Involved  
RT in EDTA Degradation from the Bacterium BNCL.";  
RL Appl. Environ. Microbiol. 67:688-695(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BNCL.  
RX MEDLINE=21091959; PubMed=11157233;  
RA Liu Y., Louie T.M., Payne J., Bohuslavsek J., Bolton H. Jr., Xun L.;  
RT "Identification, purification, and characterization of iminodiacetate  
RT oxidase from the EDTA-degrading bacterium BNCL.";  
RL Appl. Environ. Microbiol. 67:696-701(2001).  
DR EMBL; AF176564; AAG09263.1; -.  
DR HSSP; P07157; IAIIP.  
DR InterPro; IPR000795; GTP\_EFTU.  
DR PRINTS; PR00315; ELONGATNFCT.  
KW GTP-binding.  
SQ SEQUENCE 391 AA; 42794 MW; E6B41737CCD77AA6 CRC64;

Query Match 85.4%; Score 105; DB 2; Length 391;  
Best Local Similarity 90.5%; Pred. No. 8e-09;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTIGTIGHVDH 22  
| | | | | |||||:|||||||

Db 3 KGFERTKPHVNTIGTIGHVDH 23

RESULT 19

Q9RG53 ID Q9RG53 PRELIMINARY; PRT; 397 AA.  
AC Q9RG53;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE ELONGATION FACTOR TU1.  
GN TU1  
OS Streptovorticillium netropsis (Streptovorticillium flavopersicus).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
OX NCBI\_TaxID=55404;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TU1063;  
RA Oisthoorn-Tielemans L.N., Claaij N., Hilgenfeld R., Kraal B.;  
RT "Elongation factor Tu1 from the pulvomycin producer Streptomycetes  
RT mobaraense is resistant to both pulvomycin and kirromycin.";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF153618; AAF22608.1; -.  
DR HSSP; P02990; IETU  
DR InterPro; IPR000795; GTP\_EFTU.  
DR InterPro; IPR004161; GTP\_EFTU\_D2.  
DR InterPro; IPR004160; GTP\_EFTU\_D3.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
DR PRINTS; PR00315; ELONGATNFCT.  
DR PROSITE; PS00301; EFATOR.GTP; 1.  
KW GTP-binding.  
SQ SEQUENCE 397 AA; 43682 MW; 5F3AB1F3FC42914F CRC64;

Query Match 85.4%; Score 105; DB 2; Length 397;  
Best Local Similarity 85.7%; Pred. No. 8.2e-09;  
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTIGTIGHVDH 22  
| | | | | |||||:|||||||

Db 3 KAKFERTKPHVNTIGTIGHVDH 23

```
RESULT 20
Q93T39          PRELIMINARY;          PRT;    397 AA.
AC Q93T39;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN TUF1.
OS Streptomyces aureofaciens.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1894;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=84/25;
RC Kormanec J., Nguyen L.D., Novotna J., Knirschova R., Weiser J.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF368284; AAK54131.1; -.
DR SEQUENCE 397 AA; 43683 MW; BF107EA8036FEF82 CRC64;
SQ

Query Match      85.4%; Score 105; DB 2; Length 397;
Best Local Similarity 85.7%; Pred. No. 8.2e-09;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 KAKFERTKPHVNIQTIGHIDH 23

RESULT 21
Q9XD38          PRELIMINARY;          PRT;    401 AA.
AC Q9XD38;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN TUF.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=SEROVAR LAI;
RC MEDLINE=94014964; PubMed=8409911;
RA Gravekamp C., Van de Kemp H., Franzen M., Carrington D., Schoone G.J.,
RA Van Eys G.J., Everard C.O., Hartskeerl R.A., Terpstra W.J.;
RT "Detection of seven species of pathogenic leptospires by PCR using two
RT sets of primers.";
RL J. Gen. Microbiol. 139:1691-1700(1993).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=SEROVAR LAI;
RC MEDLINE=20088835; PubMed=10620683;
RA Zuerner R.L., Harskeer R.A., van de Kemp H., Bal A.E.;
RT "Characterization of the Leptospira interrogans S10-spc-alpha
RT operon.";
RL FEMS Microbiol. Lett. 182:303-308(2000).
DR EMBL; AF115283; AAD40614.1; -.
DR HSSP; P02990; 1EFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW ,GTP-binding.
```

```
SQ SEQUENCE 401 AA; 43574 MW; B8ED827C1A717BA0 CRC64;

Query Match      85.4%; Score 105; DB 2; Length 401;
Best Local Similarity 81.0%; Pred. No. 8.3e-09;
Matches 17; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 KEKFRSKPHLVNVTIGHVDH 23

RESULT 22
Q9X6G1          PRELIMINARY;          PRT;    358 AA.
AC Q9X6G1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE EF-TU (FRAGMENT).
GN TUFB.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coxiella group; Coxiella.
OX NCBI_TaxID=777;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NINE MILE PHASE I;
RC MEDLINE=20002589; PubMed=10531263;
RA Seshadri R., Hendrix L.R., Samuel J.E.;
RT "Differential expression of translational elements by life cycle
RT variants of Coxiella burnetii.";
RL Infect. Immun. 67:6026-6033(1999).
DR EMBL; AF136604; AAD32649.1; -.
DR HSSP; P02990; 1EFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW GTP-binding.
FT NON_TER 358
SQ SEQUENCE 358 AA; 39448 MW; 4355D7A33B0CCF2C CRC64;

Query Match      82.9%; Score 102; DB 2; Length 358;
Best Local Similarity 85.7%; Pred. No. 2.2e-08;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 KEKFRKPHVNIQTIGHVDH 23

RESULT 23
Q925Y6          PRELIMINARY;          PRT;    391 AA.
AC Q925Y6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PROBABLE ELONGATION FACTOR TU PROTEIN
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=1021.
RC MEDLINE=21368234; PubMed=11474104;
RX
```

RA	Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA	Batouy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA	Bourly M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA	Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA	Glox S., Godrie T., Goffeau A., Golding B., Gouzé J., Gurjal M.,
RA	Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA	Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA	Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Furnelle B.,
RA	Ramsperger U., Surzycki R., Thebaud P., Vandenbol M.,
RA	Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT	"The composite genome of the legume symbiont Sinorhizobium meliloti.";
RL	Science 293:668-672(2001).
DR	EMBL; AL591787; CAC45933.1; -.
DR	DBL; AL591787; CAC45918.1; -.
KW	Complete proteome.
SQ	SEQUENCE 391 AA; 42749 MW; 8846A3AF44753693 CRC64;
Query Match	82.9%; Score 102; DB 16; Length 391;
Best Local Similarity	85.7%; Pred. No. 2.5e-08;
Matches 18; Conservative	0; Mismatches 3; Indels 0; Gaps 0;
Qy	2 KEKENRKPHVNICITIGHVDH 22 
Dd	3 KSKFERNKPHVNICITIGHVDH 23
RESULT 24	
Q9SEF8	PRELIMINARY; PRT; 467 AA.
ID AC Q9SEF8	
DT DT 01-MAY-2000 (TrEMBLrel. 13, Created)	
DT DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE DE CHLOROPLAST TRANSLATIONAL ELONGATION FACTOR TU.	
GN GN TUFA.	
OS OS Oryza sativa (Rice).	
OC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC OC Ehrhartoidae; Oryzeae; Oryza.	
OX NCBI_TaxID=4530;	
RN RN [1]	
RP RP SEQUENCE FROM N.A.	
RC RC STRAIN=CV. JAPONICA NIPPONBARE;	
RX RX MEDLINE=20062172; PubMed=10597036;	
RA RA Lee J.H.; Lee J.W., Chung Y.Y., Paek K.H., Shin J.S., Yun C.H.,	
RA Kim J.K.;	
RT RT "Cloning and characterization of the chloroplast elongation factor EF-	
RT RT tu cDNA of Oryza sativa L.;"	
RL RL Mol. Cells 9:484-490(1999).	
DR DR EMBL; AF145053; AAF15312.1; -.	
DR DR HSSP; P07157; IAIP.	
DR DR InterPro; IPR000795; GTP_EFTU.	
DR DR InterPro; IPR004161; GTP_EFTU_D2.	
DR DR InterPro; IPR004160; GTP_EFTU_D3.	
DR DR Pfam; PF00009; GTP_EFTU; 1.	
DR DR Pfam; PF03144; GTP_EFTU_D2; 1.	
DR DR Pfam; PF03143; GTP_EFTU_D3; 1.	
DR DR PRINTS; PR00315; ELONGATNEFT.	
DR DR PROSITE; PS00301; EFATOR_GTP; 1.	
KW KW Elongation factor; GTP-binding; Protein biosynthesis.	
SQ SEQUENCE 467 AA; 50354 MW; 79A119279AF1847 CRC64;	

Query Match	82.9%	Score 102;	DB 10;	Length 467;
Best Local Similarity	94.7%	Pred. No. 3e-08;		
Matches 18; Conservative		0; Mismatches	1; Indels	0; Gaps
0;				
Qy	4	KFNRTPKPHVNIGTIGHVDH	22	
Db	63	KFERTKPHVNIGTIGHVDH	81	

```

RESULT 25
Q9RAJ9 PRELIMINARY; PRT; 393 AA.
ID AC Q9RAJ9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN TUFA.
OS Moraxella sp. TACII25.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae; Moraxella.
OX NCBI_TaxID=103760;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TAC II 25;
RA MASULLO M., Arcari P., de Paola B., Parmeggiani A., Bocchini V.;
RX MEDLINE=20563866; PubMed=11112539;
RT "Psychrophilic elongation factor Tu from the antarctic Moraxella sp.
FT Tac II 25: biochemical characterization and cloning of the encoding
RT gene.";
RL Biochemistry 39:15531-15539(2000).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=TAC II 25;
RC STRAIN=TAC II 25;
RA Arcari P.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249258; CAB65285.2; -.
HSSP; P02990; 1EFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF031144; GTP_EFTU_D2; 1.
DR Pfam; PF031143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW GTP-binding.
SQ SEQUENCE 393 AA; 43171 MW; 6A1217B78105A1B1 CRC64;

Query Match 81.3%; Score 100; DB 2; Length 393;
Best Local Similarity 81.0%; Pred. No. 5.3e-08;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 2 KEKFNRTKPHVNVGTIGHVDH 22
Db 3 KAKTERVKPHVNVGTIGHVDH 23

RESULT 26
Q9CEI0 PRELIMINARY; PRT; 395 AA.
ID AC Q9CEI0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN TUF.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OX Lactococcus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RA MEDLINE=21235186; PubMed=1137471;
RX Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RT Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -; SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL; AE006416; AAK05959.1; -.
DR HSSP; P02990; 1EFU.

```



DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Chloroplast; Elongation factor; GTP-binding; Protein biosynthesis.  
 SQ SEQUENCE 409 AA; 44707 MW; D42945D21604A57B CRC64;

Query Match 78.9%; Score 97; DB 8; Length 409;  
 Best Local Similarity 81.0%; Pred. NO. 1.7e-07;  
 Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 KEKFKPHVNIPTIGHVDH 22  
 : || | ||||| ||||| |||||  
 Db 3 RAKFKPKPHVNIPTIGHVDH 23

RESULT 30  
 Q9C2P1 PRELIMINARY; PRT; 437 AA.  
 ID Q9C2P1 AC Q9C2P1;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE PROBABLE TRANSLATION ELONGATION FACTOR EF-TU PRECURSOR,  
 DE MITOCHONDRIAL.  
 GN 17E5.70.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL513467; CAC28833.1; -  
 DR HSSP; P02990; 1EFU.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; GTP-binding; Protein biosynthesis.  
 SQ SEQUENCE 437 AA; 47848 MW; 2F64E769FE7CEE35 CRC64;

Query Match 78.9%; Score 97; DB 3; Length 437;  
 Best Local Similarity 94.4%; Pred. NO. 1.8e-07;  
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 ENRTKPHVNIPTIGHVDH 22  
 : | ||||| ||||| |||||  
 Db 41 FORTKPHVNIPTIGHVDH 58

RESULT 31  
 Q40450 PRELIMINARY; PRT; 457 AA.  
 ID Q40450 AC Q40450;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE CHLOROPLAST ELONGATION FACTOR TUA (EF-TUA) PRECURSOR (FRAGMENT).  
 GN TUA.  
 OS Nicotiana sylvestris (Wood tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LEAF;  
 RA MEDLINE=93363910; PubMed=8358028;  
 RA Murayama Y., Matsubayashi T., Sugita M., Sugiyama M.;  
 RT "Purification of chloroplast elongation factor Tu and cDNA analysis in  
 RT tobacco: the existence of two chloroplast elongation factor Tu  
 RT species.";  
 RL Plant Mol. Biol. 22:767-774(1993).  
 DR EMBL; D11375; BAA01974.1; -  
 DR HSSP; P02990; 1EFU.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; GTP-binding; Protein biosynthesis; Transit peptide.  
 FT TRANSIT 1 70 POTENTIAL.  
 FT CHAIN 71 >457 CHLOROPLAST ELONGATION FACTOR TUA  
 FT NON\_TER 457 457  
 SQ SEQUENCE 457 AA; 49732 MW; 8A30C50DC137F755 CRC64;

Query Match 78.0%; Score 96; DB 10; Length 457;  
 Best Local Similarity 89.5%; Pred. NO. 2.8e-07;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 KENRTKPHVNIPTIGHVDH 22  
 : || | ||||| ||||| |||||  
 Db 74 KEFKKPHVNIPTIGHVDH 92

RESULT 32  
 Q9AXU2 PRELIMINARY; PRT; 474 AA.  
 ID Q9AXU2 AC Q9AXU2;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE CHLOROPLAST TRANSLATIONAL ELONGATION FACTOR TU.  
 GN TUA.  
 OS Pelargonium graveolens (rose geranium).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Geraniales; Geraniaceae; Pelargonium.  
 OX NCBI\_TaxID=73200;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20550852; PubMed=11101151;  
 RA Kang C.J., Lee M.G., Cho Y.S., Lee J.W., Kyung Y.J., Shin J.S.,  
 RA Kim E.S., Kim J.K.;  
 RT "Characterization of geranium (Pelargonium graveolens) chloroplast EF-  
 RT Tu cDNA.";  
 RL Mol. Cells 10:579-583(2000).  
 DR EMBL; AF234537; AAK08141.1; -  
 DR HSSP; P02990; 1EFU.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; GTP-binding; Protein biosynthesis.  
 SQ SEQUENCE 474 AA; 51315 MW; A321742468553B65 CRC64;





```
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CG6050 PROTEIN.
GN EFTUM OR CG6050.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin C.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Murphy L., Muzny D.M., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Nixon K., Nusskern D.R., Pacle J.M.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kimons I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003818; AAF58366.1; -.
DR HSSP: P02990; 1FTU.
DR Flybase: FBgn024556; EFTUM.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR004161; GTP_EFTU_D2.
DR InterPro: IPR004160; GTP_EFTU_D3.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR Pfam: PF03143; GTP_EFTU_D3; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
KW GTP-binding; Protein biosynthesis.
SQ SEQUENCE 489 AA; 53992 MW; 9D09C36CAE82EE1 CRC64;
```

Query Match 77.2%; Score 95; DB 5; Length 489;  
Best Local Similarity 76.2%; Pred. No. 4.4e-07;  
Matches 16; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQIGHVDH 22

Db 73 KKVFTKPHCNVGTIGHVDH 93

```
RESULT 37
O78381 PRELIMINARY; PRT; 119 AA.
AC O78381;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ELONGATION FACTOR TU (FRAGMENT).
GN TUF.
OS Theileria annulata.
OG Chloroplast.
OC Eukaryota; Alveolata; Apicomplexa; Piroplasmida; Theileriidae;
OC Theileria.
OX NCBI_TaxID=5874;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HISSAR;
RA Denny P., Preiser P., Williamson D., Wilson I.;
RT "Evidence for a single origin of the 35 kb plastid DNA in
RT Apicomplexans.";
RL Protist 149:51-59(1998).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
DR EMBL: Y11429; CAA72237.1; -.
DR HSSP: P02990; 1FTU.
DR InterPro: IPR000795; GTP_EFTU.
DR Pfam: PF00009; GTP_EFTU; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
KW Chloroplast; Elongation factor; GTP-binding; Protein biosynthesis.
FT NON_TER 119
SQ SEQUENCE 119 AA; 13312 MW; 122F230BC0340BEC CRC64;

Query Match 76.4%; Score 94; DB 8; Length 119;
Best Local Similarity 71.4%; Pred. No. 1.3e-07;
Matches 15; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KQQLNKPHIINIGTIGHIDH 23

RESULT 38
O9P909 PRELIMINARY; PRT; 396 AA.
AC O9P909;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ELONGATION FACTOR TU.
GN XF2640 AND XF2628.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:43:02 ; Search time 26.88 Seconds  
(without alignments)  
78.645 Million cell updates/sec

Title: US-09-488-737-1

Perfect score: 123

Sequence: 1 MKEKENRTPHYNIGTIGHVDH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	95.9	399	2 E64670	translation elonga
2	112	91.1	394	2 B60663	translation elonga
3	112	91.1	397	2 H7285	elongation factor
4	111	90.2	394	1 EFECT	translation elonga
5	111	90.2	394	1 EFECTA	translation elonga
6	111	90.2	394	2 S13561	translation elonga
7	111	90.2	394	2 S13560	translation elonga
8	111	90.2	394	2 E64078	translation elonga
9	111	90.2	394	2 F91152	protein chain elon
10	111	90.2	394	2 G31241	protein chain elon
11	111	90.2	394	2 D82332	translation elonga
12	111	90.2	394	2 G82337	translation elonga
13	111	90.2	394	2 B6089	translation elonga
14	111	90.2	394	2 AB0457	hypothetical prote
15	111	90.2	394	2 AE0025	elongation factor
16	111	90.2	394	2 B85998	protein chain elon
17	111	90.2	394	2 AD0934	elongation factor
18	111	90.2	394	2 AD1005	elongation factor
19	109	88.6	395	2 AD1406	translation elonga
20	109	88.6	395	2 AD1782	translation elonga
21	109	88.6	409	2 S78248	translation elonga
22	108	87.8	399	2 A81392	translation elonga
23	107	87.0	391	2 AE3346	protein translatio
24	107	87.0	394	2 S78139	translation elonga
25	107	87.0	394	2 T10168	translation elonga
26	107	87.0	394	2 D81234	translation elonga
27	107	87.0	394	2 A81235	translation elonga
28	107	87.0	396	2 S55281	translation elonga
29	107	87.0	396	2 A87403	translation elonga

30	107	87.0	400	2 G72243	translation elonga
31	107	87.0	406	2 AH3344	protein translatio
32	106	86.2	394	2 C71672	translation elonga
33	106	86.2	394	2 S62726	translation elonga
34	106	86.2	394	2 H97825	elongation factor
35	106	86.2	396	2 A44795	translation elonga
36	106	86.2	396	2 S31151	translation elonga
37	106	86.2	396	2 G87143	elongation factor
38	106	86.2	397	2 F83111	elongation factor
39	106	86.2	398	2 F06663	translation elonga
40	106	86.2	398	2 D95173	translation elonga
41	106	86.2	398	2 F98039	translation elonga
42	105	85.4	394	1 EFYMTS	translation elonga
43	105	85.4	397	2 S23908	translation elonga
44	105	85.4	397	2 S23909	translation elonga
45	105	85.4	397	2 S0138	translation elonga
46	105	85.4	397	2 PC4060	translation elonga
47	105	85.4	409	1 EFEGT	translation elonga
48	104	84.6	394	1 EFYMTG	translation elonga
49	104	84.6	394	2 S73503	translation elonga
50	104	84.6	396	2 D60663	translation elonga
51	104	84.6	399	2 E71844	translation elonga
52	104	84.6	423	2 C84991	elongation factor
53	103	83.7	405	2 B70300	translation elonga
54	103	83.7	405	2 G70465	translation elonga
55	103	83.7	409	2 T07239	translation elonga
56	103	83.7	409	2 AB2348	translation elonga
57	102	82.9	391	2 C97596	elongation factor
58	102	82.9	391	2 D97594	elongation factor
59	102	82.9	391	2 AB2816	elongation factor
60	102	82.9	391	2 AC2818	elongation factor
61	102	82.9	397	2 JG5385	translation elonga
62	101	82.1	397	2 I40216	translation elonga
63	101	82.1	401	2 C70159	translation elonga
64	101	82.1	408	2 S62725	translation elonga
65	100	81.3	46	2 S23851	translation elonga
66	100	81.3	394	2 G71528	probable translati
67	100	81.3	394	2 B86500	elongation factor
68	100	81.3	394	2 F72123	translation elonga
69	100	81.3	395	2 I40828	translation elonga
70	100	81.3	399	2 F81685	translation elonga
71	99	80.5	394	2 E60663	translation elonga
72	99	80.5	395	2 E86857	translation elonga
73	99	80.5	396	1 A54536	translation elonga
74	99	80.5	405	2 E75533	translation elonga
75	99	80.5	409	1 ERITT	translation elonga
76	99	80.5	409	2 S04430	translation elonga
77	98	79.7	396	2 D6956	translation elonga
78	98	79.7	405	2 C60663	translation elonga
79	98	79.7	409	2 S73208	translation elonga
80	98	79.7	410	2 S04391	translation elonga
81	97	78.9	394	2 F89822	translational elon
82	97	78.9	396	2 A60663	translation elonga
83	97	78.9	396	2 T44381	translation elonga
84	97	78.9	406	1 S00229	translation elonga
85	97	78.9	406	1 S17146	translation elonga
86	97	78.9	406	1 S29293	translation elonga
87	97	78.9	418	2 S09153	translation elonga
88	97	78.9	495	2 E71354	probable translati
89	96	78.0	476	2 S09152	translation elonga
90	96	78.0	478	2 JQ2240	translation elonga
91	96	78.0	478	2 S36183	translation elonga
92	96	78.0	479	2 S21567	translation elonga
93	96	78.0	479	2 S60659	translation elonga
94	96	78.0	485	2 S36184	translation elonga
95	95	77.2	394	2 JCI420	translation elonga
96	95	77.2	397	2 JH0416	translation elonga
97	95	77.2	399	2 S75862	translation elonga
98	95	77.2	409	1 EFYMT	translation elonga
99	94	76.4	396	2 A82532	translation elonga
100	94	76.4	401	2 S57466	probable translati

ALIGNMENTS

RESULT 1  
E64670  
translation elongation factor EF-Tu Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 02-Feb-2001  
C:Accession: E64670  
R:Tomb, J.F.; White, O.; Kervatage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: E64670  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-399 <TOM>  
A:Cross-references: GB:AE000626; GB:AE000511; NID:G2314360; PIDN:AAD08250.1; PID:G231436  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 95.9%; Score 118; DB 2; Length 399;  
Best Local Similarity 100.0%; Pred. No. 3.2e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 KEKENRTKPHVNIPTIGHVDH 22  
|||||  
DB 3 KEKENRTKPHVNIPTIGHVDH 23  
RESULT 2  
B60663  
translation elongation factor EF-Tu Bacteroides fragilis  
C:Species: Bacteroides fragilis  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 02-Feb-2001  
C:Accession: B60663  
R:Ludwig, W.; Weizenecker, M.; Betzel, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenh  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongat  
A:Reference number: A60663; MUID:90240875  
A:Accession: B60663  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-394 <LUD>  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 91.1%; Score 112; DB 2; Length 394;  
Best Local Similarity 95.2%; Pred. No. 2.6e-09;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KEKENRTKPHVNIPTIGHVDH 22  
|||||  
DB 3 KEKENRTKPHVNIPTIGHVDH 23

RESULT 3  
H97285

elongation Factor Tu (Ef-Tu) [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C:Accession: H97285  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L  
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: H97285  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <KUR>  
A:Cross-references: GB:AB001437; PIDN:AAK81075.1; PID:G15026203; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3136  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 91.1%; Score 112; DB 2; Length 397;  
Best Local Similarity 95.2%; Pred. No. 2.6e-09;  
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 2 KEKENRTKPHVNIPTIGHVDH 22  
|||||  
DB 3 KEKENRTKPHVNIPTIGHVDH 23

RESULT 4  
EFECT  
translation elongation factor EF-Tu.B [validated] - Escherichia coli  
N:Contains: GTPase (EC 3.6.1.-)  
C:Species: Escherichia coli  
C:Date: 30-Nov-1980 #sequence\_revision 15-Oct-1982 #text\_change 19-Jan-2001  
C:Accession: A91478; A91095; A92332; I58035; G65204; A03518; A91475  
R:An, G.; Friesen, J.D.  
Gene 12, 33-39, 1980  
A:Title: The nucleotide sequence of tufB and four nearby trna structural genes of Esc  
A:Reference number: A91478; MUID:81165558  
A:Accession: A91478  
A:Molecule type: DNA  
A:Residues: 1-394 <ANG>  
A:Cross-references: GB:J01717; NID:G147968; PIDN:AAA24669.1; PID:G147969  
R:Jones, M.D.; Petersen, T.E.; Nielsen, K.M.; Magnusson, S.; Sottrup-Jensen, L.; Gaus  
Eur. J. Biochem. 108, 507-526, 1980  
A:Title: The complete amino-acid sequence of elongation factor Tu from Escherichia co  
A:Reference number: A91095; MUID:81003875  
A:Accession: A91095  
A:Molecule type: protein  
A:Residues: 2-394 <JON>  
R:Laursen, R.A.; L'Italien, J.J.; Nagarkatti, S.; Miller, D.L.  
J. Biol. Chem. 256, 8102-8109, 1981  
A:Title: The amino acid sequence of elongation factor Tu of Escherichia coli. The com  
A:Reference number: A92332; MUID:81264196  
A:Accession: A92332  
A:Molecule type: protein  
A:Residues: 2-394 <LAU>  
R:Hudson, L.; Rossi, J.; Landy, A.  
Nature 294, 422-427, 1981  
A:Title: Dual function transcripts specifying trna and mrna.  
A:Reference number: I58035; MUID:82080657  
A:Accession: I58035  
A:Status: preliminary; translated from GB/EMBL/DBDJB  
A:Molecule type: DNA  
A:Residues: 1-13 <RES>  
A:Cross-references: EMBL:X04181; NID:G43228; PIDN:CAA27777.1; PID:G43229  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G65204

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-394 <BLAT>

A:Cross-references: GB:AE000472; GB:U00096; NID:q2367333; PIDN:AAC76954.1; PID:gl790412;  
 A:Experimental source: strain K-12, substrain MG1655

R:Kawashima, T.; Berthet-colominas, C.; Wulff, M.; Cusack, S.; Leberman, R.  
 submitted to the Brookhaven Protein Data Bank, July 1996

A:Reference number: A65506; PDB:1EFU

A:Contents: annotation; X-ray crystallography, 2.5 angstroms

R:Kawashima, T.; Colominas, C.B.; Wulff, M.; Cusack, S.; Leberman, R.

Nature 379, 511-518, 1996

A:Title: The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5angstroms resol  
 A:Reference number: A58848; MUID:96170031

A:Contents: annotation; X-ray crystallography, 2.5 angstroms

C:Genetics:

A:Gene: tufB

A:Map position: 90 min

A:Start codon: GTG

C:Function:

A:Description: elongation factor Tu promotes the binding of aminoacyl-tRNA to ribosomes

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homol

C:Keywords: acetylated amino end; GTP binding; hydrolase; methylated amino acid; nucleot

F:2-394/Product: translation elongation factor Tu #status experimental <MAT>

F:13-139/Domain: translation elongation factor Tu homology <ETU>

F:19-26/Region: nucleotide-binding motif A (P-loop)

F:174-176/Region: GTP-binding NKXD motif

F:174-176/Region: GTP-binding SAK/L motif

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

F:57/Modified site: N6-methyllysine or N6,N6-dimethyllysine (Lys) (partial) #status expe

Query Match 90.2%; Score 111; DB 1; Length 394;

Best Local Similarity 90.5%; Pred. No. 3.7e-09;

Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNTKPHVNTGTIGHVDH 22

||||| |||||||:|||||||

Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 5

EFFECTA

translation elongation factor EF-Tu.A [validated] - Escherichia coli

N:Contains: GTPase (EC 3.6.1.-)

C:Species: Escherichia coli

C:Date: 30-Nov-1980 #sequence-revision 09-Jun-1994 #text-change 19-Jan-2001

C:Accession: A91475; B92332; A61304; F65127; A03518; A91478

R:Yokota, T.; Sugisaki, H.; Takanami, M.; Kaziro, Y.

Gene 12, 25-31, 1980

A:Title: The nucleotide sequence of the cloned tufA gene of Escherichia coli.

A:Reference number: A91475; MUID:81165557

A:Accession: A91475

A:Molecule type: DNA

A:Residues: 1-394 <YOK>

A:Cross-references: GB:J01690; NID:q147889; PID:gl47897

R:Jones, M.D.; Petersen, T.E.; Nielsen, K.M.; Magnusson, S.; Sottrup-Jensen, L.; Gausing

Eur. J. Biochem. 108, 507-526, 1980

A:Title: The complete amino-acid sequence of elongation factor Tu from Escherichia coli.

A:Reference number: A91095; MUID:81003875

A:Accession: B91095

A:Molecule type: protein

A:Residues: 2-394 <JON>

R:Laursen, R.A.; L'Italien, J.J.; Nagarkatti, S.; Miller, D.L.

J. Biol. Chem. 256, 8102-8109, 1981

A:Title: The amino acid sequence of elongation factor Tu of Escherichia coli. The comple

A:Reference number: A92332; MUID:81264196

A:Accession: B92332

A:Molecule type: protein

A:Residues: 2-394 <LAU>

R:Nakamura, S.; Nakayama, N.; Takahashi, K.; Kaziro, Y.

J. Biochem. 91, 1047-1063, 1982

A:Title: Primary structure of the polypeptide chain elongation factor Tu from Escherich

A:Reference number: A61304; MUID:82189911

A:Accession: A61304

A:Molecule type: protein

A:Residues: 50-264 <NAK>

A>Note: this portion of the molecule was designated fragment B

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617

A:Accession: F65127

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-394 <BLAT>

A:Cross-references: GB:AE000410; GB:U00096; NID:q1789734; PIDN:AAC76364.1; PID:gl7897

A:Experimental source: strain K-12, substrain MG1655

R:Kawashima, T.; Berthet-colominas, C.; Wulff, M.; Cusack, S.; Leberman, R.

submitted to the Brookhaven Protein Data Bank, July 1996

A:Reference number: A65506; PDB:1EFU

A:Contents: annotation; X-ray crystallography, 2.5 angstroms

R:Kawashima, T.; Colominas, C.B.; Wulff, M.; Cusack, S.; Leberman, R.

Nature 379, 511-518, 1996

A:Title: The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5angstroms re

A:Reference number: A58848; MUID:96170031

A:Contents: annotation; X-ray crystallography, 2.5 angstroms

C:Genetics:

A:Gene: tufA; strD

A:Map position: 74 min

A:Start codon: GTG

C:Function:

A:Description: elongation factor Tu promotes the binding of aminoacyl-tRNA to ribosom

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: acetylated amino end; GTP binding; hydrolase; methylated amino acid; nucl

F:2-394/Product: translation elongation factor Tu #status experimental <MAT>

F:13-139/Domain: translation elongation factor Tu homology <ETU>

F:19-26/Region: nucleotide-binding motif A (P-loop)

F:136-139/Region: GTP-binding NKXD motif

F:174-176/Region: GTP-binding SAK/L motif

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status experimental

F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #

F:57/Modified site: N6-methyllysine or N6,N6-dimethyllysine (Lys) (partial) #status e

Query Match 90.2%; Score 111; DB 1; Length 394;

Best Local Similarity 90.5%; Pred. No. 3.7e-09;

Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNTKPHVNTGTIGHVDH 22

||||| |||||||:|||||||

Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 6

SI3561

translation elongation factor EF-Tu.B - Salmonella typhimurium

C:Species: Salmonella typhimurium

C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 02-Feb-2001

C:Accession: SI3561

R:Tuohey, T.M.F.; Thompson, S.; Gesteland, R.F.; Hughes, D.; Atkins, J.F.

Biochim. Biophys. Acta 1050, 274-278, 1990

A:Title: The role of EF-Tu and other translation components in determining translocat

A:Reference number: SI3560; MUID:91002658

A:Accession: SI3561

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-394 <Tuo>

A:Cross-references: EMBL:X55117; NID:g47947; PIDN:CAA38913.1; PID:g47948

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1990

C:Genetics:

A:Gene: tufB

C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis

F:13-139/Domain: translation elongation factor Tu homology <ETU>

F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

## RESULT 7

translation elongation factor EF-Tu.A - Salmonella typhimurium  
C:Species: Salmonella typhimurium  
C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 02-Feb-2001  
C:Accession: S13560  
R:Tuohy, T.M.F.; Thompson, S.; Gesteland, R.F.; Hughes, D.; Atkins, J.F.  
Biochim. Biophys. Acta 1050, 274-278, 1990  
A:Title: The role of EF-Tu and other translation components in determining translocation  
A:Reference number: S13560; PMID:91002658  
A:Accession: S13560  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <TUO>  
A:Cross-references: EMBL:X55116  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1990  
C:Genetics:  
A:Gene: tufA  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
F:13-139/Domain: binding; nucleotide binding; P-loop; protein biosynthesis  
F:19-26/Region: translation elongation factor Tu homology <ETU>  
F:136-139/Region: nucleotide-binding motif A (P-loop)  
F:174-176/Region: GTP-binding NKXD motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

## RESULT 8

translation elongation factor EF-Tu - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 02-Feb-2001  
C:Accession: E64078; A64083  
R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; PMID:95350630  
A:Accession: E64078  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <TIG1>  
A:Cross-references: GB:U32739; GB:L42023; NID:g1573559; PIDN:AAC22236.1; PID:g1573560; T  
A:Accession: A64083  
A:Molecule type: DNA  
A:Residues: 1-394 <TIG2>  
A:Cross-references: GB:U32746; GB:L42023; NID:g1573626; PIDN:AAC22292.1; PID:g1573634; T

C:Genetics:  
A:Note: two copies of this gene are found in the Haemophilus influenzae chromosome  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

## RESULT 9

protein chain elongation factor EF-Tu [imported] - Escherichia coli (strain O157:H7,  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: F91152  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and  
A:Reference number: A99629; PMID:21156231; PMID:11258796  
A:Accession: F91152  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA37613.1; PID:g13363663; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS4190  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

## RESULT 10

protein chain elongation factor EF-Tu [imported] - Escherichia coli (strain O157:H7,  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: G91241  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and  
A:Reference number: A99629; PMID:21156231; PMID:11258796  
A:Accession: G91241  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA38326.1; PID:g13364379; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS4903  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

```
Query Match          90.2%; Score 111; DB 2; Length 394;
Best Local Similarity 90.5%; Pred. No. 3.7e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRKPHVNIQTIGHVDH 22
    |||| |||||||:|||||||
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 11
D82332
translation elongation factor EF-Tu VC0362 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: D82332
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: D82332
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <HEI>
A:Cross-references: GB:AE004124; GB:AE003852; NID:g9654770; PIDN:AAF93535.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0362
A:Map position: 1
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match          90.2%; Score 111; DB 2; Length 394;
Best Local Similarity 90.5%; Pred. No. 3.7e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRKPHVNIQTIGHVDH 22
    |||| |||||||:|||||||
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 12
G82337
translation elongation factor EF-Tu VC0321 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82337
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: G82337
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <HEI>
A:Cross-references: GB:AE004120; GB:AE003852; NID:g9654727; PIDN:AAF93494.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0321
A:Map position: 1
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match          90.2%; Score 111; DB 2; Length 394;
Best Local Similarity 90.5%; Pred. No. 3.7e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRKPHVNIQTIGHVDH 22
    |||| |||||||:|||||||
Db 3 KEKFERTKPHVNVGTIGHVDH 23
```

```
RESULT 13
D86089
hypothetical protein tufB [imported] - Escherichia coli (strain O157:H7, substrain ED
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: D86089
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Llm, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D86089
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <STO>
A:Cross-references: GB:AE005174; NID:g12518902; PIDN:AAG59176.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: tufB
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match          90.2%; Score 111; DB 2; Length 394;
Best Local Similarity 90.5%; Pred. No. 3.7e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRKPHVNIQTIGHVDH 22
    |||| |||||||:|||||||
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 14
AB0457
elongation factor Tu [imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AB0457
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0457
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-394 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC93222.1; PID:g15981670; GSPDB:GN00175
C:Genetics:
A:Gene: tufA
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match          90.2%; Score 111; DB 2; Length 394;
Best Local Similarity 90.5%; Pred. No. 3.7e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRKPHVNIQTIGHVDH 22
    |||| |||||||:|||||||
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 15
AE0025
elongation factor Tu [imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
C:Accession: AE0025
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel
```

Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AE0025

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC89064.1; PID:g15978304; GSPDB:GN00175  
C:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 16

B85998  
protein chain elongation factor EF-Tu [imported] - Escherichia coli (strain O157:H7, sub  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001  
C:Accession: B85998  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: AB8480; MUID:21074935; PMID:11206551  
A:Accession: B85998  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <STO>  
A:Cross-references: GB:AE005174; NID:g12517957; PIDN:ARG58446.1; GSPDB:GN00145; UWGP:Z46  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 17

AD0934  
elongation factor Tu [imported] - Salmonella enterica subsp. enterica serovar Typhi (str  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD0934  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AD0934  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD09494.1; PID:g16504611; GSPDB:GN00176  
C:Genetics:

A:Gene: tufb  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom  
Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 18

AD1005  
elongation factor Tu [imported] - Salmonella enterica subsp. enterica serovar Typhi (C  
C:Species: Salmonella enterica subsp. enterica serovar Typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AD1005  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AD1005  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD08168.1; PID:g16505144; GSPDB:GN00176  
C:Genetics:  
A:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 90.2%; Score 111; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 3.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| |||||||:|||||||  
Db 3 KEKFERTKPHVNVGTIGHVDH 23

RESULT 19

AD1406  
translation elongation factor EF-Tu homolog tufa [imported] - Listeria monocytogenes  
C:Species: Listeria monocytogenes  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AD1406  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.;  
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1406  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-395 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAD00866.1; PID:g16412153; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 88.6%; Score 109; DB 2; Length 395;  
Best Local Similarity 90.5%; Pred. No. 7.4e-09;



Matches 19; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
:||||:|||||  
Db 3 KEKFDKSKPHVNIQTIGHVDH 23

RESULT 20  
AD1782  
translation elongation factor EF-Tu homolog tufa [imported] - Listeria innocua (strain C  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AD1782  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma  
ok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,  
A.; Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1782  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-395 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC98028.1; PID:g16415338; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: tufa  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

Query Match 88.6%; Score 109; DB 2; Length 395;  
Best Local Similarity 90.5%; Pred. No. 7.4e-09;  
Matches 19; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
:||||:|||||  
Db 3 KEKFDKSKPHVNIQTIGHVDH 23

RESULT 21  
S78248  
translation elongation factor EF-Tu - Odontella sinensis chloroplast  
C:Species: chloroplast Odontella sinensis  
C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 02-Feb-2001  
C:Accession: S78248  
R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sinensis  
A:Reference number: S78238  
A:Accession: S78248  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-409 <ROW>  
A:Cross-references: EMBL:Z67753; NID:g1185127; PIDN:CAA91621.1; PID:g1185138  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
C:Genetics:  
A:Gene: tufa  
A:Gene: chloroplast  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo  
C:Keywords: chloroplast; GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 88.6%; Score 109; DB 2; Length 409;  
Best Local Similarity 90.5%; Pred. No. 7.7e-09;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
:||||:|||||  
Db 3 KEKFNRTKPHVNIQTIGHVDH 23

RESULT 22  
AB1392  
translation elongation factor EF-Tu Cj0470 [similarity] - Campylobacter jejuni (strai  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Sep-2000  
C:Accession: AB1392  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chil  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals  
A:Reference number: AB1250; MUID:20150912  
A:Accession: AB1392  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-399 <PAR>  
A:Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB75108.1; PID:g696  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: tuf; Cj0470  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom  
F:13-139/Domain: translation elongation factor Tu homology <ETU>

Query Match 87.8%; Score 108; DB 2; Length 399;  
Best Local Similarity 90.5%; Pred. No. 1.1e-08;  
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
:||||:|||||  
Db 3 KEKFNRTKPHVNIQTIGHVDH 23

RESULT 23  
AE3346  
protein translation elongation factor Tu (EF-tu) [imported] - Brucella melitensis (st  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
R:DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov  
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AE3346  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-391 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51936.1; PID:g17982694; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0755  
A:Map position: I  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 87.0%; Score 107; DB 2; Length 391;  
Best Local Similarity 90.5%; Pred. No. 1.5e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
:||||:|||||  
Db 3 KSKFNRTKPHVNIQTIGHVDH 23

RESULT 24  
S78139  
translation elongation factor EF-Tu.A - Reclinomonas americana (ATCC 50394) mitochond  
N:Alternate names: translation elongation factor tufa

C:Species: mitochondrion *Reclinomonas americana*  
A:Variety: ATCC 50394  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 02-Feb-2001  
C:Accession: S78139  
R:Lang, B.F.; Burger, G.; O'Kelly, C.J.; Cedergren, R.; Golding, G.B.; Lemieux, C.; Sank  
Nature 387, 493-497, 1997  
A:Title: An ancestral mitochondrial DNA resembling a eubacterial genome in miniature.  
A:Reference number: S78127; MUID:97311393  
A:Accession: S78139  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-394 <LAN>  
A:Cross-references: EMBL:AF007261; NID:g22583325; PIDN:AA011872.1; PID:g2258338  
A:Experimental source: ATCC 50394  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1997  
C:Genetics:  
A:Gene: tufA  
A:Genome: mitochondrion  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; methylated amino acid; mitochondrion; nucleotide binding; P-loop  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Thr, Asn, Lys, Asp, Ser) #sta  
F:57/Modified site: N6-methyllysine or N6,N6-dimethyllysine (Lys) #status Predicted

Query Match 87.0%; Score 107; DB 2; Length 394;  
Best Local Similarity 90.5%; Pred. No. 1.5e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| ||||| ||||| ||||| |||||  
Db 3 KEKFERTKPHCNIGTIGHVDH 23

RESULT 25  
T10168  
translation elongation factor Tu - *Neisseria gonorrhoeae*  
C:Species: *Neisseria gonorrhoeae*  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jun-2000  
C:Accession: T10168  
R:Porcella, S.F.; Belland, R.J.; Judd, R.C.  
Microbiology 142, 2481-2489, 1996  
A:Title: Identification of an EF-Tu protein that is periplasm-associated and processed i  
A:Reference number: Z16970  
A:Accession: T10168  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-394 <POR>  
A:Cross-references: EMBL:L36380; NID:g790857; PID:g790859  
A:Experimental source: strain MS1  
C:Genetics:  
A:Gene: tufA  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match 87.0%; Score 107; DB 2; Length 394;  
Best Local Similarity 85.7%; Pred. No. 1.5e-08;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| ||||| ||||| ||||| |||||  
Db 3 KEKFERSKPHVNVGTIGHVDH 23

RESULT 26  
DB1234  
translation elongation factor Tu NMB0139 [imported] - *Neisseria meningitidis* (strain MC5  
C:Species: *Neisseria meningitidis*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: DB1234

R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: AB1000; MUID:20175755  
A:Accession: DB1234  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <TET>  
A:Cross-references: GB:AE002372; GB:AE002098; NID:g7225350; PIDN:AAF40598.1; PID:g722  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB0139  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom

Query Match 87.0%; Score 107; DB 2; Length 394;  
Best Local Similarity 85.7%; Pred. No. 1.5e-08;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||||| ||||| ||||| ||||| |||||  
Db 3 KEKFERSKPHVNVGTIGHVDH 23

RESULT 27  
AB1235  
translation elongation factor Tu NMB0124 [imported] - *Neisseria meningitidis* (strain  
C:Species: *Neisseria meningitidis*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
C:Accession: AB1235; DB2006; CB2008  
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: AB1000; MUID:20175755  
A:Accession: AB1235  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <TET>  
A:Cross-references: GB:AE002371; GB:AE002098; NID:g7225337; PIDN:AAF40583.1; PID:g722  
A:Experimental source: serogroup B, strain MC58  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491  
A:Reference number: AB1775; MUID:20222556  
A:Accession: DB2006  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <PAR>  
A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83449.1; PID:g737  
A:Experimental source: serogroup A, strain Z2491  
A:Accession: CB2008  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-394 <PA2>  
A:Cross-references: GB:AL162752; GB:AL157959; NID:g7378778; PIDN:CAB83464.1; PID:g737  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics: <PAR1>  
A:Gene: tufA1; NMB0124; NMA0134  
C:Genetics: <PAR2>  
A:Gene: tufA2; NMA0149  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom  
F:13-139/Domain: translation elongation factor Tu homology <ETU>

Query Match 87.0%; Score 107; DB 2; Length 394;  
Best Local Similarity 85.7%; Pred. No. 1.5e-08;

Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNICTIGHVDH 22  
|||||:|||||:|||||:|||||

Db 3 KEKERSKPHVNVGTIGHVDH 23

RESULT 28  
S55281  
translation elongation factor EF-Tu - Stigmatella aurantiaca  
C:Species: Stigmatella aurantiaca  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 02-Feb-2001  
C:Accession: S55281; S52237  
R:Bremaud, L.; Fremaux, C.; Laalami, S.; Cenatiempo, Y.  
Nucleic Acids Res. 23, 1737-1743, 1995  
A:Title: Genetic and molecular analysis of the trna-tufB operon of the myxobacterium Stigmatella aurantiaca  
A:Reference number: S55281; MUID:95303606  
A:Accession: S55281  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-396 <BRE>  
A:Cross-references: EMBL:X82820; NID:g609255; PIDN:CAA58029.1; PID:g609256  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1994  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 87.0%; Score 107; DB 2; Length 396;  
Best Local Similarity 90.5%; Pred. No. 1.5e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNICTIGHVDH 22  
|||||:|||||:|||||:|||||

Db 3 KEKERNKPHVNICTIGHVDH 23

RESULT 29  
A87403  
translation elongation factor EF-Tu [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 18-Jul-2001  
C:Accession: A87403; E87645  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
n, J.; Leub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87403  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-396 <STO>  
A:Cross-references: GB:AE005673; NID:g13422569; PIDN:AAK23221.1; GSPDB:GN00148  
A:Accession: E87645  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-396 <ST2>  
A:Cross-references: GB:AE005673; NID:g13424877; PIDN:AAK25161.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC1240; CC3199  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog

Query Match 87.0%; Score 107; DB 2; Length 396;  
Best Local Similarity 90.5%; Pred. No. 1.5e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNICTIGHVDH 22

Db 3 KEKERTKPHCNIGTIGHVDH 23

RESULT 30  
G72243  
translation elongation factor EF-Tu - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 02-Feb-2001  
C:Accession: G72243; A48314  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MUID:99287316  
A:Accession: G72243  
A:Molecule type: DNA  
A:Residues: 1-400 <ARN>  
A:Cross-references: GB:AE001799; GB:AE000512; NID:g4982067; PIDN:AAD36569.1; PID:g498  
A:Experimental source: strain MSB8  
R:Bachleitner, M.; Ludwig, W.; Stetter, K.O.; Schleifer, K.H.  
FEMS Microbiol. Lett. 57, 115-120, 1989  
A:Title: Nucleotide sequence of the gene coding for the elongation factor Tu from the  
A:Reference number: A48314  
A:Accession: A48314  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-39, 'V', 41-400 <BAC>  
A:Cross-references: GB:M27479; NID:g154836; PIDN:AAA27415.1; PID:g552037  
C:Genetics:  
A:Gene: TMI502  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-139/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:136-139/Region: GTP-binding NKXD motif  
F:174-176/Region: GTP-binding SAK/L motif  
F:25,26,62,136,137,139,174/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

Query Match 87.0%; Score 107; DB 2; Length 400;  
Best Local Similarity 85.7%; Pred. No. 1.5e-08;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNICTIGHVDH 22  
|||||:|||||:|||||:|||||

Db 3 KEKFRTRKPHVNVGTIGHIDH 23

RESULT 31  
AH3344  
protein translation elongation factor Tu (Ef-Tu) [imported] - Brucella melitensis (st  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C:Accession: AH3344  
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov  
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AH3344  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-406 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51923.1; PID:g17982679; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0742  
A:Map position: I  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu hom



A:Gene: tuf  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-141/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (p-loop)  
F:138-141/Region: nucleotide-binding motif A (p-loop)  
F:175-177/Region: GTP-binding NKXD motif  
F:25,26,64,138,139,141,175/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 86.2%; Score 106; DB 2; Length 396;  
Best Local Similarity 90.5%; Pred. No. 2.1e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
| | | | | | | | | | | | | | | | | | | | | |  
DB 3 KAKFERTKPHVNIQTIGHVDH 23  
| | | | | | | | | | | | | | | | | | | | | |

RESULT 36

S31151  
translation elongation factor EF-Tu - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 02-Feb-2001  
C:Accession: S31151; S34954; J2262  
R:Honore, N.; Bergh, S.; Chanteau, S.; Doucet-Populaire, F.; Eigmeier, K.; Garnier, T.;  
del, P.R.; Sittisombut, N.; Wu-Hunter, S.; Cole, S.T.  
Mol. Microbiol. 7, 207-214, 1993  
A:Title: Nucleotide sequence of the first cosmid from the Mycobacterium leprae genome ph  
A:Reference number: S31143; MUID:93188701  
A:Accession: S31151  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-396 <HON>  
A:Cross-references: EMBL:L13216; NID:g44382; PIDN:CAA78674.1; PID:g581338  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992  
R:Silbaq, F.; Bercovier, H.  
Nucleic Acids Res. 21, 3327, 1993  
A:Title: Nucleotide sequence of Mycobacterium leprae elongation factor (EF-Tu) gene.  
A:Reference number: S34954; MUID:93341950  
A:Accession: S34954  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-128, 'GVPIYLVNLSDAVDDELLV', 153-280, 'L', 282-348, 'G', 350-396 <SIL>  
A:Cross-references: EMBL:L13216; NID:g293241; PIDN:AAA71969.1; PID:g293242  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1993  
R:Dhandayuthapani, S.; Banu, M.J.; Kashiwabara, Y.  
J. Biochem. 115, 664-669, 1994  
A:Title: Cloning and sequence determination of the gene coding for the elongation factor  
A:Reference number: JC2262; MUID:94375410  
A:Accession: JC2262  
A:Molecule type: DNA  
A:Residues: 'V', 2-128, 'GVPIYILVA', 137, 'NKSDAVDDEELLE', 151, 'V', 153-280, 'L', 282-348, 'G', 350-396  
A:Cross-references: DDBJ:D13869  
C:Genetics:  
A:Gene: tuf  
A:Start codon: GTG  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolog  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-141/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (p-loop)  
F:83-87/Region: GTP binding #status predicted  
F:138-141/Region: GTP binding #status predicted

Query Match 86.2%; Score 106; DB 2; Length 396;  
Best Local Similarity 90.5%; Pred. No. 2.1e-08;  
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNIQTIGHVDH 22  
| | | | | | | | | | | | | | | | | | | | | |  
DB 3 KAKFERTKPHVNIQTIGHVDH 23  
| | | | | | | | | | | | | | | | | | | | | |

translation elongation factor EF-Tu - Streptococcus oralis  
C:Species: Streptococcus oralis  
C:Date: 28-Apr-1993 #sequence\_revision 28-Apr-1993 #text\_change 02-Feb-2001  
C:Accession: F60663  
R:Ludwig, W.; Weizenegger, M.; Betzl, D.; Leidel, E.; Lenz, T.; Ludvigsen, A.; Moellenh  
Arch. Microbiol. 153, 241-247, 1990  
A:Title: Complete nucleotide sequences of seven eubacterial genes coding for the elongat  
A:Reference number: A60663; MUID:90240875  
A:Accession: F60663  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA

A:Residues: 1-398 <LUD>  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo  
C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis  
F:13-142/Domain: translation elongation factor Tu homology <ETU>  
F:19-26/Region: nucleotide-binding motif A (P-loop)  
F:139-142/Region: GTP-binding NKXD motif  
F:177-179/Region: GTP-binding SAK/L motif  
F:25,26,65,139,140,142,177/Binding site: Mg-GTP (Lys, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 86.2%; Score 106; DB 2; Length 398;  
Best Local Similarity 85.7%; Pred. No. 2.le-08;  
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
|||.:|:|||||:|||||  
Db 3 KEYDRSKPHVNIQTIGHVDH 23

RESULT 40

D95173  
translation elongation factor Tu [imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001  
C:Accession: D95173  
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Held  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: D95173  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-398 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK75581.1; PID:g14972978; GSPDB:GN00164; TIGR:SP4  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SP1489  
C:Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo

Query Match 86.2%; Score 106; DB 2; Length 398;  
Best Local Similarity 85.7%; Pred. No. 2.le-08;  
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
|||.:|:|||||:|||||  
Db 3 KEYDRSKPHVNIQTIGHVDH 23

Search completed: August 22, 2002, 07:46:12  
Job time: 190 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:43:01 ; Search time 21.52 Seconds  
(without alignments)  
24.970 Million cell updates/sec

Title: US-09-488-737-1

Perfect score: 123

Sequence: 1 MKRKNRTKPHVNICTIGHVDH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

- Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	111	90.2	394	US-09-140-466-6	Sequence 6, Appli
2	102	82.9	408	US-09-140-466-4	Sequence 4, Appli
3	99	80.5	409	US-09-140-466-3	Sequence 3, Appli
4	95	77.2	409	US-09-140-466-5	Sequence 5, Appli
5	78	63.4	410	US-09-140-466-2	Sequence 2, Appli
6	51	41.5	462	US-08-299-351-1	Sequence 1, Appli
7	51	41.5	462	US-08-371-377-18	Sequence 18, Appli
8	51	41.5	462	5225348-1	Patent No. 5225348
9	50	40.7	460	US-09-174-768-4	Sequence 4, Appli
10	49	39.8	688	US-08-221-817-19	Sequence 19, Appli
11	49	39.8	688	US-08-454-439-19	Sequence 19, Appli
12	49	39.8	688	PCT-US94-10487-19	Sequence 19, Appli
13	49	39.8	689	US-08-221-817-18	Sequence 18, Appli
14	49	39.8	689	US-08-454-439-18	Sequence 18, Appli
15	49	39.8	689	PCT-US94-10487-18	Sequence 18, Appli
16	49	39.8	699	US-08-221-817-20	Sequence 20, Appli
17	49	39.8	699	US-08-454-439-20	Sequence 20, Appli
18	49	39.8	699	PCT-US94-10487-20	Sequence 20, Appli
19	48	39.0	499	US-08-820-170A-40	Sequence 40, Appli
20	48	39.0	499	US-09-055-699-40	Sequence 40, Appli
21	48	39.0	499	US-09-273-565-40	Sequence 40, Appli
22	48	39.0	499	US-09-565-538-40	Sequence 40, Appli
23	45	36.6	480	US-08-272-255-12	Sequence 12, Appli
24	45	36.6	480	PCT-US95-08565-12	Sequence 12, Appli
25	42	34.1	620	US-09-442-100-11	Sequence 11, Appli
26	42	34.1	1142	US-09-106-075A-89	Sequence 89, Appli
27	42	34.1	1148	US-09-106-075A-90	Sequence 90, Appli

28	41	33.3	1	US-08-626-994A-3	Sequence 3, Appli
29	41	33.3	339	US-08-957-742-3	Sequence 3, Appli
30	41	33.3	364	US-08-626-994A-1	Sequence 1, Appli
31	41	33.3	364	US-08-957-742-1	Sequence 1, Appli
32	40.5	32.9	321	US-09-039-609-4	Sequence 4, Appli
33	40.5	32.9	334	US-08-646-981-16	Sequence 16, Appli
34	40	32.5	336	US-09-131-648-1	Sequence 1, Appli
35	40	32.5	407	US-08-989-370-5	Sequence 5, Appli
36	40	32.5	407	US-08-989-370-6	Sequence 6, Appli
37	40	32.5	455	US-08-889-841B-46	Sequence 46, Appli
38	40	32.5	474	US-08-889-841B-36	Sequence 36, Appli
39	40	32.5	474	US-08-889-841B-39	Sequence 39, Appli
40	40	32.5	508	US-08-472-240A-16	Sequence 16, Appli
41	40	32.5	608	US-09-257-490-13	Sequence 13, Appli
42	40	32.5	701	US-09-132-028-2	Sequence 2, Appli
43	40	32.5	850	US-08-448-603A-28	Sequence 28, Appli
44	40	32.5	850	US-09-134-075-28	Sequence 28, Appli
45	40	32.5	855	US-07-956-483-14	Sequence 14, Appli
46	40	32.5	872	US-07-956-483-12	Sequence 12, Appli
47	40	32.5	887	US-08-472-240A-5	Sequence 5, Appli
48	40	32.5	906	US-08-472-240A-3	Sequence 3, Appli
49	39	31.7	91	US-08-325-253-10	Sequence 10, Appli
50	39	31.7	93	US-08-817-946-5	Sequence 5, Appli
51	39	31.7	171	US-09-129-030-56	Sequence 56, Appli
52	39	31.7	221	US-08-698-978-2	Sequence 2, Appli
53	39	31.7	221	US-08-470-953A-3	Sequence 3, Appli
54	39	31.7	247	US-08-501-126-19	Sequence 19, Appli
55	39	31.7	248	US-09-189-060B-2	Sequence 2, Appli
56	39	31.7	248	US-09-189-060B-12	Sequence 12, Appli
57	39	31.7	248	US-08-470-953A-6	Sequence 6, Appli
58	39	31.7	420	US-08-845-998-8	Sequence 8, Appli
59	39	31.7	420	US-09-206-537-8	Sequence 8, Appli
60	39	31.7	420	US-09-430-854-8	Sequence 8, Appli
61	39	31.7	436	US-09-150-213-4	Sequence 4, Appli
62	39	31.7	1140	US-08-657-641-7	Sequence 7, Appli
63	39	31.7	1140	PCT-US94-07233-7	Sequence 7, Appli
64	38.5	31.3	194	US-08-118-469A-7	Sequence 7, Appli
65	38.5	31.3	194	US-08-909-119-7	Sequence 7, Appli
66	38.5	31.3	491	US-09-314-268-3	Sequence 3, Appli
67	38.5	31.3	1507	5268270-2	Patent No. 5268270
68	38	30.9	31	US-07-978-692-3	Sequence 3, Appli
69	38	30.9	200	US-08-275-526C-24	Sequence 24, Appli
70	38	30.9	227	US-08-275-526C-31	Sequence 31, Appli
71	38	30.9	278	US-08-392-828C-4	Sequence 4, Appli
72	38	30.9	278	US-09-330-945-8	Sequence 8, Appli
73	38	30.9	483	US-08-889-841B-31	Sequence 31, Appli
74	38	30.9	485	US-08-889-841B-8	Sequence 8, Appli
75	38	30.9	487	US-08-889-841B-33	Sequence 33, Appli
76	38	30.9	498	US-08-889-841B-5	Sequence 5, Appli
77	38	30.9	498	US-08-889-841B-5	Sequence 5, Appli
78	38	30.9	519	US-08-472-240A-18	Sequence 18, Appli
79	38	30.9	865	US-07-956-483-13	Sequence 13, Appli
80	38	30.9	887	US-08-472-240A-4	Sequence 4, Appli
81	38	30.9	915	US-08-480-917-2	Sequence 2, Appli
82	38	30.9	915	US-09-138-736-2	Sequence 2, Appli
83	38	30.9	1264	US-07-789-915A-6	Sequence 6, Appli
84	38	30.9	1264	US-08-005-002C-6	Sequence 6, Appli
85	38	30.9	1264	US-08-487-203A-6	Sequence 6, Appli
86	37	30.1	144	US-09-001-944-12	Sequence 12, Appli
87	37	30.1	144	US-09-240-004A-12	Sequence 12, Appli
88	37	30.1	145	US-09-001-944-10	Sequence 10, Appli
89	37	30.1	145	US-09-240-004A-10	Sequence 10, Appli
90	37	30.1	149	US-08-765-381-7	Sequence 7, Appli
91	37	30.1	166	US-09-001-944-2	Sequence 2, Appli
92	37	30.1	166	US-09-001-944-4	Sequence 4, Appli
93	37	30.1	166	US-09-240-004A-2	Sequence 2, Appli
94	37	30.1	166	US-09-240-004A-4	Sequence 4, Appli
95	37	30.1	189	US-08-865-297-2	Sequence 2, Appli
96	37	30.1	330	US-09-267-031-16	Sequence 16, Appli
97	37	30.1	363	US-07-772-032-2	Sequence 2, Appli
98	37	30.1	363	US-09-157-077-9	Sequence 9, Appli
99	37	30.1	363	PCT-US92-08012-2	Sequence 2, Appli
100	37	30.1	528	US-08-889-841B-44	Sequence 44, Appli

*David J. Lawrence*







5225348-1  
 ; Patent No. 5225348  
 ; APPLICANT: HAGATA, SHIGEKAZU; SUGANO, SUMIO; KIM, DONG W.;  
 ; UETSUKI, TAICHI; KAZIRO, YOSHITO  
 ; TITLE OF INVENTION: DNA FRAGMENT AND EXPRESSION PLASMID  
 ; CONTAINING THE DNA FRAGMENT  
 ; NUMBER OF SEQUENCES: 9  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/447,823  
 ; FILING DATE: 08-DEC-1989  
 ; SEQ ID NO:1:  
 ; LENGTH: 462  
 5225348-1

Query Match 41.5%; Score 51; DB 6; Length 462;  
 Best Local Similarity 60.0%; Pred. No. 1.9;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 RTKPHVNICTIGHVD 21  
 : | | | | | | | | | |  
 Db 3 KEKTHNIVVIGHVD 17

RESULT 9  
 US-09-174-768-4  
 ; Sequence 4, Application US/09174768  
 ; Patent No. 6265185  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Muller, Sven  
 ; APPLICANT: Dalbauge, Henrik  
 ; TITLE OF INVENTION: No. 6265185el Yeast Promoters Suitable For Expression Cloning In  
 ; TITLE OF INVENTION: Heterologous Expression Of Proteins In Yeast  
 ; FILE REFERENCE: 4791.204-US  
 ; CURRENT APPLICATION NUMBER: US/09/174,768  
 ; CURRENT FILING DATE: 1998-10-19  
 ; EARLIER APPLICATION NUMBER: 0589/96  
 ; EARLIER FILING DATE: 1996-05-21  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 460  
 ; TYPE: PRT  
 ; ORGANISM: EF-lalpha  
 US-09-174-768-4

Query Match 40.7%; Score 50; DB 4; Length 460;  
 Best Local Similarity 60.0%; Pred. No. 2.8;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 RTKPHVNICTIGHVD 21  
 : | | | | | | | | | |  
 Db 3 KEKTHNIVVIGHVD 17

RESULT 10  
 US-08-221-817-19  
 ; Sequence 19, Application US/08221817  
 ; Patent No. 5532151  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chantry, David  
 ; APPLICANT: Gray, Patrick W.  
 ; APPLICANT: Hoekstra, Merle F.  
 ; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor  
 ; TITLE OF INVENTION: Kinase GRK6  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; ADDRESS: Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois

COUNTRY: USA  
 ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/221,817  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/123,932  
 ; FILING DATE: 17 SEP 1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: NO. 5532151land, Greta E.  
 ; REGISTRATION NUMBER: 35,302  
 ; REFERENCE/DOCKET NUMBER: 31981  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (312) 474-6300  
 ; TELEFAX: (312) 474-0448  
 ; TELEX: 25-3856  
 ; INFORMATION FOR SEQ ID NO: 19:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 688 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-221-817-19

Query Match 39.8%; Score 49; DB 1; Length 688;  
 Best Local Similarity 43.8%; Pred. No. 6.3;  
 Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 FNRTKPHVNICTIGHV 20  
 : | | | | | | | | | |  
 Db 342 FSKKKPHASVGTGYM 357

RESULT 11  
 US-08-454-439-19  
 ; Sequence 19, Application US/08454439  
 ; Patent No. 5591618  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chantry, David  
 ; APPLICANT: Gray, Patrick W.  
 ; APPLICANT: Hoekstra, Merle F.  
 ; TITLE OF INVENTION: A No. 5591618el G Protein-Coupled Receptor  
 ; TITLE OF INVENTION: Kinase GRK6  
 ; NUMBER OF SEQUENCES: 24  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; ADDRESS: Borun  
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: USA  
 ; ZIP: 60606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/454,439  
 ; FILING DATE: 30-MAY-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/221,817  
 ; FILING DATE: 31-MAR-1994  
 ; APPLICATION NUMBER: 08/123,932  
 ; FILING DATE: 17 SEP 1993

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: NO. 5591618and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-454-439-19

Query Match 39.8%; Score 49; DB 1; Length 688;
Best Local Similarity 43.8%; Pred. No. 6.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 5 FNRTKPHVNIQTIGHV 20
Db 342 FSKKKPHASVGTGYM 357

RESULT 12
PCT-US94-10487-19
; Sequence 19, Application PC/TUS94/10487
; GENERAL INFORMATION:
; APPLICANT: ICOS Corporation
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/221,817
; FILING DATE: 31 MAR 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
PCT-US94-10487-19

Query Match 39.8%; Score 49; DB 5; Length 688;
Best Local Similarity 43.8%; Pred. No. 6.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 5 FNRTKPHVNIQTIGHV 20
Db 342 FSKKKPHASVGTGYM 357

RESULT 13
US-08-221-817-18
; Sequence 18, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5532151and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 689 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-221-817-18

Query Match 39.8%; Score 49; DB 1; Length 689;
Best Local Similarity 43.8%; Pred. No. 6.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 5 FNRTKPHVNIQTIGHV 20
Db 342 FSKKKPHASVGTGYM 357

RESULT 14
US-08-454-439-18
; Sequence 18, Application US/08454439

```

```
; Patent No. 5591618
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,439
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/221,817
; FILING DATE: 31-MAR-1994
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5591618and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 689 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-439-18

Query Match 39.8%; Score 49; DB 1; Length 689;
Best Local Similarity 43.8%; Pred. No. 6.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 FNRTKPHVNIGTIGHV 20
   1:: 111 ::11 1::
Db 342 FSKKKPHASVGTGYM 357

RESULT 15
PCT-US94-10487-18
; Sequence 18, Application PC/TUS9410487
; GENERAL INFORMATION:
; APPLICANT: ICOS Corporation
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/221,817
; FILING DATE: 31 MAR 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 689 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-10487-18

Query Match 39.8%; Score 49; DB 5; Length 689;
Best Local Similarity 43.8%; Pred. No. 6.3;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 FNRTKPHVNIGTIGHV 20
   1:: 111 ::11 1::
Db 342 FSKKKPHASVGTGYM 357

RESULT 16
US-08-221-817-20
; Sequence 20, Application US/08221817
; Patent No. 5532151
; GENERAL INFORMATION:
; APPLICANT: Chantry, David
; APPLICANT: Gray, Patrick W.
; APPLICANT: Hoekstra, Merle F.
; TITLE OF INVENTION: A No. 5532151el G Protein-Coupled Receptor
; TITLE OF INVENTION: Kinase GRK6
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/221,817
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-454-439-20

Query Match 39.8%; Score 49; DB 1; Length 699;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 FNRTKPHVNIQTIGHV 20
Db 343 FSKKKPHASVGTGHM 358
    |:: ||| ::|| |::
    |:: ||| ::|| |::

RESULT 18
PCT-US94-10487-20
; Sequence 20, Application PC/TUS9410487
; GENERAL INFORMATION:
; APPLICANT: ICOS Corporation
; TITLE OF INVENTION: A Novel G Protein-Coupled Receptor
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10487
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/221,817
; FILING DATE: 31 MAR 1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/123,932
; FILING DATE: 17 SEP 1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31981
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-10487-20

Query Match 39.8%; Score 49; DB 5; Length 699;
Best Local Similarity 43.8%; Pred. No. 6.4;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 5 FNRTKPHVNIQTIGHV 20
Db 343 FSKKKPHASVGTGHM 358
    |:: ||| ::|| |::
    |:: ||| ::|| |::

```

```
RESULT 19
US-08-820-170A-40
; Sequence 40, Application US/08820170A
; Patent No. 5831058
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,170A
; FILING DATE:
; CLASSIFICATION: 536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-820-170A-40

Query Match 39.0%; Score 48; DB 2; Length 499;
Best Local Similarity 69.2%; Pred. No. 6.3;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 9 KPHVNICTIGHVD 21
| | | | |
Db 72 KEHVNVVFIGHVD 84

RESULT 20
US-09-055-699-40
; Sequence 40, Application US/09055699
; Patent No. 6005088
; GENERAL INFORMATION:
; APPLICANT: Tsutomu, FUJIWARA
; APPLICANT: Takeshi, WATANABE
; APPLICANT: Masato, HORIE
; APPLICANT: Toyomasa, KATAGIRI
; TITLE OF INVENTION: HUMAN GENE
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/09/055,699
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/820,170
; FILING DATE:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 293-7060
; TELEFAX: (202) 293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-055-699-40

Query Match 39.0%; Score 48; DB 3; Length 499;
Best Local Similarity 69.2%; Pred. No. 6.3;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 9 KPHVNICTIGHVD 21
| | | | |
Db 72 KEHVNVVFIGHVD 84

RESULT 21
US-09-273-565-40
; Sequence 40, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-273-565-40

Query Match 39.0%; Score 48; DB 4; Length 499;
Best Local Similarity 69.2%; Pred. No. 6.3;
Matches 9; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 9 KPHVNICTIGHVD 21
| | | | |
Db 72 KEHVNVVFIGHVD 84

RESULT 22
US-09-565-538-40
; Sequence 40, Application US/09565538
; Patent No. 633404
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
```

```

; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/565,538
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 09/273,565
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 09/055,699
; PRIOR FILING DATE: 1998-04-07
; PRIOR APPLICATION NUMBER: 08/820,170
; PRIOR FILING DATE: 1997-03-19
; PRIOR APPLICATION NUMBER: JP 63410/1996
; PRIOR FILING DATE: 1996-03-19
; PRIOR APPLICATION NUMBER: JP 69163/1997
; PRIOR FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-565-538-40

```

```

Query Match          39.08; Score 48; DB 4; Length 499;
Best Local Similarity 69.28; Pred. No. 6.3;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Qy 9 KPHVNIQIGHVD 21
   | | | | |
Db 72 KEVNVVFIQGHV 84

```

```

RESULT 23
US-08-272-255-12
; Sequence 12, Application US/08272255
; Patent No. 5824859
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272,255
; FILING DATE: 08-JUL-1994
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: OPN-1795
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-272-255-12

Query Match          36.68; Score 45; DB 2; Length 480;
Best Local Similarity 47.18; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQIG 18
   : | : | | |
Db 237 BEPTSRLSPHLKFGTIG 253

RESULT 24
PCT-US95-08565-12
; Sequence 12, Application PC/TUS9508565
; GENERAL INFORMATION:
; APPLICANT: Cashmore, Anthony R.
; APPLICANT: Ahmad, Margaret
; APPLICANT: Lin, Chentao
; TITLE OF INVENTION: Blue Light Photoreceptors and Methods of
; TITLE OF INVENTION: Using the Same
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/08565
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/272,255
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: OPN-1795
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-08565-12

Query Match          36.68; Score 45; DB 5; Length 480;
Best Local Similarity 47.18; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQIG 18
   : | : | | |
Db 237 BEPTSRLSPHLKFGTIG 253

RESULT 25
US-09-442-100-11
; Sequence 11, Application US/09442100

```

; Patent No. 6359193  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Tian  
; APPLICANT: Tao, Wufan  
; APPLICANT: Wang, Weiyl  
; APPLICANT: Zhang, Sheng  
; APPLICANT: Yu, Wan  
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS  
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/442,100  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/411,111  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 6523-003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 620 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-442-100-11

Query Match 34.1%; Score 42; DB 4; Length 620;  
Best Local Similarity 54.5%; Pred. No. 74;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 NRTKPHVNICT 16  
| :||:|:|  
Db 3 NTRPHLNLGT 13

RESULT 26  
US-09-106-075A-89  
; Sequence 89, Application US/09106075A  
; Patent No. 6316250  
; GENERAL INFORMATION:  
; APPLICANT: Hjelle MD, Brian  
; APPLICANT: Jenison, Steve  
; TITLE OF INVENTION: Molecular Clones Producing Recombinant DNA Antigens of  
; TITLE OF INVENTION: the HARDS Virus.  
; FILE REFERENCE: 10312-8U1, Hjelle et al. (210312.0009)  
; CURRENT APPLICATION NUMBER: US/09/106,075A  
; CURRENT FILING DATE: 1998-06-29  
; PRIOR APPLICATION NUMBER: 08/210,762  
; PRIOR FILING DATE: 1994-03-22  
; PRIOR APPLICATION NUMBER: 08/141,035  
; PRIOR FILING DATE: 1993-10-26  
; PRIOR APPLICATION NUMBER: 08/120,096  
; PRIOR FILING DATE: 1993-09-13

; PRIOR APPLICATION NUMBER: 08/111,519  
; PRIOR FILING DATE: 1993-08-25  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 89  
; LENGTH: 1142  
; TYPE: PRT  
; ORGANISM: Prospekt Hill virus  
US-09-106-075A-89

Query Match 34.1%; Score 42; DB 4; Length 1142;  
Best Local Similarity 36.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNICTIGHVD 21  
:||:|:|:|:|  
Db 930 QSFNITEPHITSNSLEWVD 948

RESULT 27  
US-09-106-075A-90  
; Sequence 90, Application US/09106075A  
; Patent No. 6316250  
; GENERAL INFORMATION:  
; APPLICANT: Hjelle MD, Brian  
; APPLICANT: Jenison, Steve  
; TITLE OF INVENTION: Molecular Clones Producing Recombinant DNA Antigens of  
; TITLE OF INVENTION: the HARDS Virus.  
; FILE REFERENCE: 10312-8U1, Hjelle et al. (210312.0009)  
; CURRENT APPLICATION NUMBER: US/09/106,075A  
; CURRENT FILING DATE: 1998-06-29  
; PRIOR APPLICATION NUMBER: 08/210,762  
; PRIOR FILING DATE: 1994-03-22  
; PRIOR APPLICATION NUMBER: 08/141,035  
; PRIOR FILING DATE: 1993-10-26  
; PRIOR APPLICATION NUMBER: 08/120,096  
; PRIOR FILING DATE: 1993-09-13  
; PRIOR APPLICATION NUMBER: 08/111,519  
; PRIOR FILING DATE: 1993-08-25  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 90  
; LENGTH: 1148  
; TYPE: PRT  
; ORGANISM: Puumala virus  
US-09-106-075A-90

Query Match 34.1%; Score 42; DB 4; Length 1148;  
Best Local Similarity 31.6%; Pred. No. 1.5e+02;  
Matches 6; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNICTIGHVD 21  
:||:|:|:|:|  
Db 934 QSFNITEPHITSALEWID 952

RESULT 28  
US-08-626-994A-3  
; Sequence 3, Application US/08626994A  
; Patent No. 5798244  
; GENERAL INFORMATION:  
; APPLICANT: Shuichi TSUJI et al.  
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAc \ 2,8-  
; TITLE OF INVENTION: SIALYLTRANSFERASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenblum & Bernstein, P.L.C.  
; STREET: 1941 Roland Clarke Place  
; CITY: Reston  
; STATE: Virginia  
; COUNTRY: U.S.A.



```

; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/626,994A
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-626-994A-3

Query Match 33.3%; Score 41; DB 1; Length 339;
Best Local Similarity 40.0%; Pred. No. 54;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHVDH 22
Db 71 KNFSLTSSVRIGQLMHYDY 90

RESULT 29
US-08-957-742-3
; Sequence 3, Application US/08957742
; Patent No. 6017743
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAC ' 2,8-
; TITLE OF INVENTION: STALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-626-994A-1

Query Match 33.3%; Score 41; DB 1; Length 364;
Best Local Similarity 40.0%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHVDH 22
Db 96 KNFSLTSSVRIGQLMHYDY 115

TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-957-742-3

Query Match 33.3%; Score 41; DB 3; Length 339;
Best Local Similarity 40.0%; Pred. No. 54;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHVDH 22
Db 71 KNFSLTSSVRIGQLMHYDY 90

RESULT 30
US-08-626-994A-1
; Sequence 1, Application US/08626994A
; Patent No. 5798244
; GENERAL INFORMATION:
; APPLICANT: Shuichi TSUJI et al.
; TITLE OF INVENTION: Sia' 2,3Gala 1,4GlcNAC ' 2,8-
; TITLE OF INVENTION: STALYLTRANSFERASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenblum & Bernstein, P.L.C.
; STREET: 1941 Roland Clarke Place
; CITY: Reston
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 20191
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Word Perfect 5.1+ (ASCII)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/626,994A
; FILING DATE: April 3, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 77469/1995
; FILING DATE: April 3, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold Turk
; REGISTRATION NUMBER: 33,094
; REFERENCE/DOCKET NUMBER: P14595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 716-1191
; TELEFAX: (703) 716-1180
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; US-08-626-994A-1

Query Match 33.3%; Score 41; DB 1; Length 364;
Best Local Similarity 40.0%; Pred. No. 59;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHVDH 22
Db 96 KNFSLTSSVRIGQLMHYDY 115

TELECOMMUNICATION INFORMATION:

```

```

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/039,609
; FILING DATE: 16-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/060,623
; FILING DATE: 01-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestija, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GP-70275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 321 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-039-609-4

Query Match          32.9%; Score 40.5; DB 3; Length 321;
Best Local Similarity 40.9%; Pred. No. 62;
Matches 9; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY      1 MKEKFNRTKPHVNIIGTIGHVDH 22
        :||:   : : || | | 1: |||

Db       279 VKELPQERPAVN-QTVAEYDH 299

RESULT 33
US-08-646-981-16
; Sequence 16, Application US/08646981
; Patent No. 5852183
; GENERAL INFORMATION:
; APPLICANT: MAEDA, HIROAKI
; APPLICANT: EDÅ, YASUYUKI
; APPLICANT: KIMACHI, KAZUHIKO
; APPLICANT: ONO, YOICHI
; APPLICANT: TOKIYOSHI, SACHIO
; TITLE OF INVENTION: DOG-MOUSE HETEROHYBRIDOMA AND GENE
; TITLE OF INVENTION: FRAGMENT CODING FOR CONSTANT REGION OF CANINE
; TITLE OF INVENTION: IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,981
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WEINER, MARC S
; REGISTRATION NUMBER: 32,181
; REFERENCE/DOCKET NUMBER: 1488-106
; INFORMATION FOR SEQ ID NO: 16:

```

SEQUENCE CHARACTERISTICS:  
LENGTH: 334 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-646-981-16

Query Match 32.5%; Score 40; DB 2; Length 334;  
Best Local Similarity 45.0%; Pred. No. 77;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 2 KEFNRTKPHVNIQTIGHV 21  
:|||||:|:|  
Db 178 EEQFNQTYRVSVLPIGHQD 197

RESULT 34  
US-09-131-648-1  
Sequence 1, Application US/09131648  
Patent No. 6158920  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Guegler, Karl J.  
APPLICANT: Patterson, Chandra  
TITLE OF INVENTION: EXTRACELLULAR ADHESIVE PROTEINS  
FILE REFERENCE: PF-0576 US  
CURRENT APPLICATION NUMBER: US/09/131,648  
CURRENT FILING DATE: 1998-08-10  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: PERL Program  
SEQ ID NO 1  
LENGTH: 336  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE: -  
OTHER INFORMATION: 2635136  
US-09-131-648-1

Query Match 32.5%; Score 40; DB 4; Length 336;  
Best Local Similarity 50.0%; Pred. No. 78;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 4 KFNRTKPHVNIQTI 17  
:|||||:|  
Db 99 RPHTTKPHVICNTL 112

RESULT 35  
US-08-989-370-5  
Sequence 5, Application US/08989370  
Patent No. 6013268  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,370  
FILING DATE: 12-DEC-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.404C5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 407 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-989-370-5

Query Match 32.5%; Score 40; DB 3; Length 407;  
Best Local Similarity 38.9%; Pred. No. 96;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHV 20  
:|:|:|:|  
Db 146 QKLOAEAPHVIVGTGGRV 163

RESULT 36  
US-08-989-370-6  
Sequence 6, Application US/08989370  
Patent No. 6013268  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/989,370  
FILING DATE: 12-DEC-1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.404C5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 407 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-989-370-6

Query Match 32.5%; Score 40; DB 3; Length 407;  
Best Local Similarity 38.9%; Pred. No. 96;  
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 3 EKFNRTKPHVNIQTIGHV 20

Db 146 OKLQAEAPHIVVGTGPRV 163  
:| ||: || | |

## RESULT 37

US-08-889-841B-46

; Sequence 46, Application US/08889841B

; GENERAL INFORMATION:

; APPLICANT: Berman, Phillip W.

; TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE

; FILE REFERENCE: 14918-703CIP

; CURRENT APPLICATION NUMBER: US/08/889,841B

; PRIOR FILING DATE: 1997-07-08

; PRIOR APPLICATION NUMBER: US 60/676,737

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 46

; LENGTH: 455

; TYPE: PRT

; ORGANISM: HIV

US-08-889-841B-46

Query Match 32.5%; Score 40; DB 3; Length 455;

Best Local Similarity 46.7%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKFNRTKPHVNIGTI 17

: || | | | |

Db 189 KKFNGTGCTNVSTV 203

## RESULT 38

US-08-889-841B-36

; Sequence 36, Application US/08889841B

; GENERAL INFORMATION:

; APPLICANT: Berman, Phillip W.

; TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE

; FILE REFERENCE: 14918-703CIP

; CURRENT APPLICATION NUMBER: US/08/889,841B

; PRIOR FILING DATE: 1997-07-08

; PRIOR APPLICATION NUMBER: US 60/676,737

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 36

; LENGTH: 474

; TYPE: PRT

; ORGANISM: HIV

US-08-889-841B-36

Query Match 32.5%; Score 40; DB 3; Length 474;

Best Local Similarity 46.7%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKFNRTKPHVNIGTI 17

: || | | | |

Db 186 KKFNGTGCTNVSTV 200

## RESULT 39

US-08-889-841B-39

; Sequence 39, Application US/08889841B

; GENERAL INFORMATION:

; APPLICANT: Berman, Phillip W.

; TITLE OF INVENTION: HIV ENVELOPE POLYPEPTIDES AND VACCINE

; FILE REFERENCE: 14918-703CIP

; CURRENT APPLICATION NUMBER: US/08/889,841B

; PRIOR FILING DATE: 1997-07-08

; PRIOR APPLICATION NUMBER: US 60/676,737

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 39

; LENGTH: 474

; TYPE: PRT

; ORGANISM: HIV

US-08-889-841B-39

; NUMBER OF SEQ ID NOS: 57

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 39

; LENGTH: 474

; TYPE: PRT

; ORGANISM: HIV

US-08-889-841B-39

Query Match 32.5%; Score 40; DB 3; Length 474;

Best Local Similarity 46.7%; Pred. No. 1.1e+02;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKFNRTKPHVNIGTI 17

: || | | | |

Db 186 KKFNGTGCTNVSTV 200

## RESULT 40

US-08-472-240A-16

; Sequence 16, Application US/08472240A

; Patent No. 6284248

; GENERAL INFORMATION:

; APPLICANT: KIENY, Marie-Paule

; TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

; TITLE OF INVENTION: gp160 VARIANT

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BURNS, DOANE, SWECKER &amp; MATHIS

; STREET: P.O. Box 1404

; CITY: Alexandria

; STATE: Virginia

; COUNTRY: United States

; ZIP: 22313-1404

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,240A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/956,483

; FILING DATE: 31-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Teskin, Robin L.

; REGISTRATION NUMBER: 35,030

; REFERENCE/DOCKET NUMBER: 017753-055

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 836-6620

; TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 508 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-472-240A-16

Query Match 32.5%; Score 40; DB 4; Length 508;

Best Local Similarity 46.7%; Pred. No. 1.2e+02;

Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 EKFNRTKPHVNIGTI 17

: || | | | |

Db 232 KKFNGTGCTNVSTV 246

Search completed: August 22, 2002, 07:44:32





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:43:01 ; Search time 51.87 Seconds  
(Without alignments)  
47.111 Million cell updates/sec

Title: US-09-488-737-1

Perfect score: 123

Sequence: 1 MKEKENRTKPHVNIQTIGHVDH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_032802.\*

- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	123	100.0	22	18	AAW33404
2	123	100.0	22	19	AAW73036
3	118	95.9	399	19	AAW73035
4	118	95.9	399	22	AAU35822
5	111	90.2	394	21	AAU35890
6	111	90.2	394	22	AAU34769
7	111	90.2	394	22	AAU34838
8	111	90.2	394	22	AAU35466
9	111	90.2	394	22	AAU35476
10	111	90.2	409	22	AAU38371
11	106	86.2	396	22	AAU383183
N-terminal sequenc					
Helicobacter pylori					
Helicobacter pylori					
Helicobacter pylori					
E. coli proliferat					
E. coli cellular p					
E. coli cellular p					
Haemophilus influe					
Haemophilus influe					
Salmonella typhi c					
C glutamicum prote					

12	106	86.2	397	22	AAU36402	Pseudomonas aerugi
13	106	86.2	397	22	AAU36409	Pseudomonas aerugi
14	106	86.2	398	22	AAU37578	Streptococcus pneu
15	105	85.4	396	13	AAU20243	Sequence of transl
16	103	83.7	396	18	AAW30303	Elongation factor
17	102	82.9	396	13	AAU20242	Sequence of transl
18	102	82.9	396	13	AAU20245	Sequence of elfamy
19	102	82.9	396	13	AAU20246	Sequence of elfamy
20	102	82.9	396	13	AAU20247	Sequence of elfamy
21	102	82.9	396	13	AAU20248	Sequence of elfamy
22	102	82.9	465	21	AAG40781	Zea mays protein f
23	101	82.1	395	22	AAU33391	Enterococcus faeca
24	101	82.1	395	22	AAU34889	Enterococcus faeca
25	100	81.3	394	22	AAU66514	Protonibacterium
26	100	81.3	394	20	AAU34672	Chlamydia pneumonia
27	100	81.3	394	22	AAU38889	C. pneumoniae CT32
28	100	81.3	394	22	AAU38909	C. trachomatis CT3
29	100	81.3	397	20	AAU36864	Protein involved i
30	100	81.3	436	22	AAU39992	Protonibacterium
31	97	78.9	310	22	AAG82109	S. epidermidis ope
32	97	78.9	394	21	AAU96769	Staphylococcus aur
33	97	78.9	394	22	AAU34078	Staphylococcus aur
34	97	78.9	394	22	AAU36849	Staphylococcus aur
35	96	78.0	77	21	AAG12253	Zea mays protein f
36	96	78.0	86	21	AAG12252	Zea mays protein f
37	96	78.0	111	21	AAG12476	Zea mays protein f
38	96	78.0	129	21	AAG12474	Zea mays protein f
39	96	78.0	156	21	AAG44634	Zea mays protein f
40	96	78.0	175	21	AAG44633	Zea mays protein f
41	96	78.0	476	21	AAG18102	Arabidopsis thalia
42	96	78.0	476	21	AAG52762	Arabidopsis thalia
43	96	78.0	497	21	AAG18101	Arabidopsis thalia
44	96	78.0	497	21	AAG52761	Arabidopsis thalia
45	95	77.2	489	22	ABB62027	Drosophila melanog
46	94	76.4	426	21	AAU03444	Candida albicans e
47	91	74.0	397	21	AAG29847	Arabidopsis thalia
48	91	74.0	397	21	AAG34905	Arabidopsis thalia
49	91	74.0	454	21	AAG29846	Arabidopsis thalia
50	91	74.0	454	21	AAG34904	Arabidopsis thalia
51	91	74.0	455	21	AAG29845	Arabidopsis thalia
52	91	74.0	470	21	AAG34903	Arabidopsis thalia
53	89	72.4	452	22	ABB50268	Mitochondrial EF-T
54	89	72.4	479	22	AAU56937	Human prostate can
55	89	72.4	479	22	AAU23533	Novel human enzyme
56	88	71.5	388	13	AAU20244	Sequence of transl
57	78	63.4	410	22	AAU07211	Plasmodium falcipa
58	72	58.5	456	22	ABB58293	Drosophila melanog
59	66	53.7	428	22	AAU96210	Putative P. abysssi
60	62	50.4	87	21	AAU44062	Human cancer assoc
61	62	50.4	116	22	AAG75706	Human colon cancer
62	62	50.4	118	21	AAU58481	Lung cancer associ
63	62	50.4	227	22	ABG19061	Novel human diagno
64	62	50.4	411	22	AAU96700	Putative P. abysssi
65	62	50.4	470	22	ABG11316	Novel human diagno
66	62	50.4	474	22	AAU32988	Novel human secret
67	62	50.4	477	22	ABG11419	Novel human diagno
68	62	50.4	484	22	ABG19062	Novel human diagno
69	62	50.4	533	22	AAU33181	Novel human secret
70	57	46.3	61	21	AAU58508	Lung cancer associ
71	56	45.5	133	21	AAU58722	Breast and ovarian
72	56	45.5	220	22	ABG17756	Novel human diagno
73	56	45.5	492	22	ABG22109	Novel human diagno
74	56	45.5	3680	22	ABG22108	Novel human diagno
75	54.5	44.3	34	21	AAU35976	Arabidopsis thalia
76	54.5	44.3	90	21	AAU33972	Arabidopsis thalia
77	54.5	44.3	101	21	AAU35974	Arabidopsis thalia
78	54.5	44.3	101	21	AAU37127	Arabidopsis thalia
79	54.5	44.3	121	21	AAU37126	Arabidopsis thalia
80	54.5	44.3	275	21	AAU35134	Arabidopsis thalia
81	54.5	44.3	431	21	AAU23949	Arabidopsis thalia
82	54.5	44.3	449	21	AAU25187	Arabidopsis thalia
83	54.5	44.3	449	21	AAU32345	Arabidopsis thalia
84	54.5	44.3	449	21	AAU35730	Arabidopsis thalia

85 54.5 44.3 44.9 21 AAG45353 Arabidopsis thalia  
86 54.5 44.3 44.9 21 AAG45356 Arabidopsis thalia  
87 54.5 44.3 44.9 21 AAG45359 Arabidopsis thalia  
88 54.5 44.3 44.9 21 AAG47136 Arabidopsis thalia  
89 54.5 44.3 44.9 21 AAG47144 Arabidopsis thalia  
90 54.5 44.3 470 21 AAG47135 Arabidopsis thalia  
91 54 43.9 155 21 AAG12390 Zea mays protein f  
92 54 43.9 284 21 AAG23944 Arabidopsis thalia  
93 54 43.9 284 21 AAG46907 Arabidopsis thalia  
94 54 43.9 300 21 AAG23943 Arabidopsis thalia  
95 54 43.9 301 21 AAG46906 Arabidopsis thalia  
96 53 43.1 133 21 AAG12391 Zea mays protein f  
97 52 42.3 110 21 AAG03900 Human secreted pro  
98 52 42.3 458 21 AAG35256 Zea mays protein f  
99 52 42.3 464 21 AAG35255 Zea mays protein f  
100 52 42.3 465 21 AAG24689 Arabidopsis thalia

## ALIGNMENTS

RESULT 1  
AAW33404  
ID AAW33404 standard; peptide; 22 AA.  
XX  
AC AAW33404;  
XX  
DT 05-MAR-1998 (first entry)  
XX  
DE N-terminal sequence of 50 kD membrane protein of H. pylori.  
XX  
KW membrane protein; vaccine; immunisation; immunoassay;  
KW diagnostic; affinity purification; gastritis; peptic ulcer;  
KW gastric cancer.  
XX  
OS Helicobacter pylori.  
XX  
PN WO9712909-A1.  
XX  
PD 10-APR-1997.  
XX  
PF 04-OCT-1996; 96WO-FR01552.  
XX  
PR 04-OCT-1995; 95FR-0011890.  
XX  
PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.  
XX  
PI Lissolo L;  
XX  
DR WPI; 1997-226163/20.  
XX  
PT New Helicobacter pylori membrane proteins - and related  
PT polypeptide(s) and antibodies, useful for active or passive  
PT immunisation and diagnosis  
XX  
PS Claim 4; Page 30; 48pp; French.

XX New purified Helicobacter pylori proteins are obtained from a  
CC membrane fraction and have apparent molecular weights (by  
CC electrophoresis on= 10% polyacrylamide gel in presence of SDS) of 54,  
CC 50, 32-35 and 30 kD respectively, of which the 54 kD protein does not  
CC react with anti-catalase antiserum. The present sequence represents the  
CC N-terminal of the 50 kD protein. The proteins and purified peptides  
CC derived from them can be used (by active immunisation) to treat or  
CC prevent H. pylori infections (gastritis, peptic ulcers, gastric cancer  
CC etc.). Antibodies raised against the protein can be used for passive  
CC immunisation. Derived peptides and antibodies are also useful as  
CC reagents for diagnostic immunoassays, and antibodies are further useful  
CC for affinity purification of the new proteins.

XX Sequence 22 AA;

Query Match 100.0%; Score 123; DB 18; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.9e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MKEKFNRTKPHVNIQTIGHVDH 22  
DB 1 mkekfnrtkphvniqtighvdh 22  
|||||  
RESULT 2  
AAW73036  
ID AAW73036 standard; Peptide; 22 AA.  
XX  
AC AAW73036;  
XX  
DT 02-FEB-1999 (first entry)  
XX  
DE Helicobacter pylori 50 kDa polypeptide N-terminal peptide.  
XX  
KW GHPO -750; Infection; gastritis; ulcer; vaccine; diagnosis; therapy.  
XX  
OS Helicobacter pylori strain ATCC 43579.

XX WO9843479-A1.

XX 08-OCT-1998.

XX 31-MAR-1998; 98WO-US06421.

XX 01-APR-1997; 97US-0834566.

XX 01-APR-1997; 97US-0831310.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS ET VACCINS.

XX Al-Garawi A, Kleanthous H, Lissolo L, Miller C, Tomb J;

XX WPI; 1998-568251/48.

XX New isolated Helicobacter polynucleotides - used to develop products  
PT for the diagnosis, prevention and treatment of Helicobacter  
PT infections and gastroduodenal diseases

XX Example 1.C.4; Page 56; 184pp; English.

XX This is the N-terminal sequence of a 50 kDa protein, designated  
CC GHPO 750, isolated by immunoaffinity chromatography from a membrane  
CC fraction of Helicobacter pylori ATCC 43579. The invention provides  
CC Helicobacter polypeptides (see AAW73022-32 and AAW73034-35) and  
CC polynucleotides (see AAW72001, AAV07912-21 and AAV07963-64) that can be  
CC used in methods for the treatment and prevention of Helicobacter  
CC infections.

XX Sequence 22 AA;

Query Match 100.0%; Score 123; DB 19; Length 22;  
Best Local Similarity 100.0%; Pred. No. 4.9e-13;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKEKFNRTKPHVNIQTIGHVDH 22  
DB 1 mkekfnrtkphvniqtighvdh 22  
|||||

RESULT 3

AAW73035

ID AAW73035 standard; Protein; 399 AA.

XX

AC AAW73035;

XX 02-FEB-1999 (first entry)

XX



DE Helicobacter pylori 32 kDa polypeptide GHPO 750.  
 KW GHPO 750; infection; gastritis; ulcer; vaccine; diagnosis;  
 KW therapy.  
 XX Helicobacter pylori.  
 OS  
 XX WO9843479-A1.  
 PN  
 XX  
 XX PD 08-OCT-1998.  
 XX  
 XX PF 31-MAR-1998; 98WO-US06421.  
 XX  
 XX PR 01-APR-1997; 97US-0834566.  
 PR  
 PR 01-APR-1997; 97US-0831310.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS ET VACCINS.  
 PI  
 PI Al-Garawi A, Kleanthous H, Lissolo L, Miller C, Tomb J;  
 XX  
 XX WPI: 1998-568251/48.  
 DR  
 DR N-PSDB; AAV07964.

XX New isolated Helicobacter polynucleotides - used to develop products  
 PT for the diagnosis, prevention and treatment of Helicobacter  
 PT infections and gastroduodenal diseases  
 XX  
 PS Claim 9; Page 150-152; 184pp; English.  
 XX  
 CC This is the amino acid sequence of a 50 kDa Helicobacter pylori  
 CC polypeptide designated GHPO 750. It was deduced from an isolated  
 CC genomic DNA sequence (see AAV07963). The invention provides a family  
 CC of 76 kDa Helicobacter polypeptides (see AAW73022-32), as well as  
 CC GHPO 750 and a 32 kDa polypeptide (see AAW73034), and also  
 CC polynucleotides (see AAV72001, AAV07912-21 and AAV07963-64) encoding  
 CC them, expression cassettes, and methods for producing the unprocessed or  
 CC mature polypeptides in host cells. The polypeptides can be used in  
 CC vaccination methods to prevent or treat Hb infection in a mammal.  
 CC Methods and products of the invention allow treatment and  
 CC prevention of gastroduodenal diseases associated with Hb  
 CC infections, including acute, chronic, and atrophic gastritis, and  
 CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. Detection  
 CC and diagnostic methods are also provided. GHPO 750 was  
 CC demonstrated to be a protective antigen. Unlike the other  
 CC polypeptides of the invention, GHPO 750 is not a secreted protein.  
 XX  
 SQ Sequence 399 AA;

Query Match 95.9%; Score 118; DB 19; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-11;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KEKENRTKPHVNICTIGHVDH 22  
 |||||  
 Db 3 kekfnrtkphvniqthghvdh 23

RESULT 4  
 AAU35822  
 ID AAU35822 standard; Protein; 399 AA.  
 XX  
 AC AAU35822;  
 XX  
 XX 14-FEB-2002 (first entry)  
 DT  
 DE Helicobacter pylori cellular proliferation protein #135.

XX Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX  
 OS Helicobacter pylori.

XX WO200170955-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX PF 21-MAR-2001; 2001WO-US09180.  
 XX  
 XX 21-MAR-2000; 2000US-191078P.  
 PR  
 PR 23-MAY-2000; 2000US-206848P.  
 PR  
 PR 26-MAY-2000; 2000US-207727P.  
 PR  
 PR 23-OCT-2000; 2000US-242578P.  
 PR  
 PR 27-NOV-2000; 2000US-253625P.  
 PR  
 PR 22-DEC-2000; 2000US-257931P.  
 PR  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI  
 XX WPI: 2001-611495/70.  
 DR  
 DR N-PSDB; AAS53681.  
 XX  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 PT  
 XX Example 3; Seq ID No 11415; 51pp; English.  
 PS  
 XX

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 399 AA;

Query Match 95.9%; Score 118; DB 22; Length 399;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-11;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KEKENRTKPHVNICTIGHVDH 22  
 |||||  
 Db 3 kekfnrtkphvniqthghvdh 23

RESULT 5  
 AAB15890  
 ID AAB15890 standard; Protein; 394 AA.  
 XX  
 AC AAB15890;  
 XX  
 XX 05-OCT-2000 (first entry)  
 DT

XX E. coli proliferation associated protein sequence SEQ ID NO:247.  
 DE  
 XX Escherichia coli; E. coli; proliferation; inhibition; screening;  
 KW antimicrobial; bacterial growth; antisense therapy; antibacterial.  
 XX

OS Escherichia coli.  
 PN WO200044906-A2.  
 XX PD 03-AUG-2000.  
 XX 27-JAN-2000; 2000WO-US02200.  
 XX PF 27-JAN-1999; 99US-0117405.  
 XX PA (ELIT-) ELITRA PHARM INC.  
 XX Zyskind J, Ohlsen KL, Trawick J, Forsyth RA, Froelich JM, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI WPI: 2000-514822/46.  
 DR N-PSDB; AAA65894.  
 XX Novel polynucleotides and polypeptides associated with microorganism  
 PT proliferation, used to identify inhibitors of bacterial growth and  
 PT proliferation, for use in antisense therapy -  
 XX Claim 11; Page 177-178; 316pp; English.  
 CC AAA65809 to AAA65889 and AAA66058 to AAA66138 represent nucleotide  
 CC sequences derived from Escherichia coli which inhibit E. coli  
 CC proliferation. AAA65890 to AAA66055 and AAA65886 to AAA66040 represent  
 CC nucleotide and protein sequences associated with E. coli proliferation.  
 CC AAA66056 and AAA66057 represent primers used for sequencing E. coli  
 CC proliferation inhibiting nucleotide inserts in an example from the  
 CC present invention. Methods from the present invention can be used to  
 CC identify a proliferation- required gene in a microorganism, by contacting  
 CC a microorganism with a proliferation-required gene activity inhibitory  
 CC nucleic acid identified in another organism, and determining if  
 CC inhibition occurs in the second microorganism. The nucleic acid sequences  
 CC identified as being required for bacterial growth and proliferation, can  
 CC be used for antisense therapy for killing bacteria.  
 XX Sequence 394 AA;

Query Match 90.2%; Score 111; DB 21; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 1e-09;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||| |||||:|||||||  
 Db 3 kekferkphvntighvdh 23

RESULT 6  
 AAU34769  
 ID AAU34769 standard; Protein; 394 AA.  
 XX AC AAU34769;

DT 14-FEB-2002 (first entry)  
 XX E. coli cellular proliferation protein #350.  
 XX Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX Escherichia coli.  
 OS WO200170955-A2.  
 PN 27-SEP-2001.  
 XX 21-MAR-2001; 2001WO-US09180.  
 XX 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX (ELIT-) ELITRA PHARM INC.  
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI WPI: 2001-611495/70.  
 DR N-PSDB; AAS52628.  
 XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX Example 3; Seq ID No 10362; 511pp; English.  
 CC The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 394 AA;

Query Match 90.2%; Score 111; DB 22; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 1e-09;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||| |||||:|||||||  
 Db 3 kekferkphvntighvdh 23

RESULT 7  
 AAU34838  
 ID AAU34838 standard; Protein; 394 AA.  
 XX AC AAU34838;

DT 14-FEB-2002 (first entry)  
 XX E. coli cellular proliferation protein #419.  
 XX Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 XX Escherichia coli.  
 OS WO200170955-A2.  
 PN 27-SEP-2001.  
 XX 21-MAR-2001; 2001WO-US09180.  
 PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX WPI; 2001-611495/70.  
 DR N-PSDB; AAS52697.  
 XX

PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID No 10431; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 394 AA;

Query Match 90.2%; Score 111; DB 22; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 1e-09;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KEKENRTKPHVNIQIGHVDH 22  
 ||||| |||||:|||||  
 Db 3 kekferkphvngtighvdh 23

RESULT 8  
 AAU35466  
 ID AAU35466 standard; Protein; 394 AA.

AC AAU35466;

DT 14-FEB-2002 (first entry)

XX Haemophilus influenzae cellular proliferation protein #107.

DE Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.

XX Haemophilus influenzae.

OS WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX

PR 21-MAR-2000; 2000US-191078P.  
 PR 23-MAY-2000; 2000US-206848P.  
 PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX

PA (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX WPI; 2001-611495/70.  
 DR N-PSDB; AAS53325.  
 XX

PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 XX  
 PS Example 3; Seq ID No 11059; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 394 AA;

Query Match 90.2%; Score 111; DB 22; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 1e-09;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 KEKENRTKPHVNIQIGHVDH 22  
 ||||| |||||:|||||  
 Db 3 kekferkphvngtighvdh 23

RESULT 9  
 AAU35476  
 ID AAU35476 standard; Protein; 394 AA.

AC AAU35476;

DT 14-FEB-2002 (first entry)

XX Haemophilus influenzae cellular proliferation protein #117.

DE Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.

XX Haemophilus influenzae.

OS WO200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX

```

XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX PA
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR N-PSDB; AAS53335.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 11069; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 394 AA;
SQ
Query Match 90.2%; Score 111; DB 22; Length 394;
Best Local Similarity 90.5%; Pred. No. 1e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 kekferkphvngvtighvdh 23
|||||
2 kekferkphvngvtighvdh 23
RESULT 10
AAU38371
ID AAU38371 standard; Protein: 409 AA.
XX
XX AAU38371;
XX
XX 14-FEB-2002 (first entry)
XX
XX Salmonella typhi cellular proliferation protein #262.
DE
XX Antisense; prokaryotic cellular proliferation protein;
KW antibiotic; antibacterial; drug design.
XX
XX Salmonella typhi.
OS
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX

```

```

PF 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX PA
XX (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI Yamamoto RT, Xu HH;
XX
XX WPI: 2001-611495/70.
DR N-PSDB; AAS56230.
XX
XX New polynucleotides for the identification and development of
PT antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Example 3; Seq ID No 13964; 51lpp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to
CC prokaryotic cellular proliferation, their use in identifying the
CC genes, their use in the discovery of novel antibiotics, the essential
CC genes themselves and the encoded proteins. The prokaryotes used are
CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC invention is also useful for the identification of potential new targets
CC for antibiotic development. The antisense nucleic acids can also be used
CC to identify proteins used in proliferation, to express these proteins,
CC and to obtain antibodies capable of binding to the expressed proteins.
CC The proteins can be used to screen compounds in rational drug discovery
CC programmes. The antisense nucleic acid sequence is also useful to screen
CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 409 AA;
SQ
Query Match 90.2%; Score 111; DB 22; Length 409;
Best Local Similarity 90.5%; Pred. No. 1.1e-09;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 18 kekferkphvngvtighvdh 38
|||||
2 kekferkphvngvtighvdh 38
RESULT 11
AAG93183
ID AAG93183 standard; Protein: 396 AA.
XX
XX AAG93183;
XX
XX 26-SEP-2001 (first entry)
XX
XX C glutamicum protein fragment SEQ ID NO: 6937.
DE
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX
XX Corynebacterium glutamicum.
OS
XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX

```



PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.

XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX  
 DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS4268.

XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -

XX  
 PS Example 3; Seq ID No 12002; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 397 AA;

Query Match 86.2%; Score 106; DB 22; Length 397;  
 Best Local Similarity 85.7%; Pred. No. 6.7e-09;  
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||| |||||:|||||  
 Db 3 kekfernkphvntgtnvdh 23

RESULT 14

AAU37578  
 ID AAU37578 standard; Protein; 398 AA.

XX  
 AC AAU37578;

DT 14-FEB-2002 (first entry)

DE Streptococcus pneumoniae cellular proliferation protein #7.

XX Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.

XX Streptococcus pneumoniae.

OS WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.  
 PR 23-OCT-2000; 2000US-242578P.  
 PR 27-NOV-2000; 2000US-253625P.  
 PR 22-DEC-2000; 2000US-257931P.  
 PR 16-FEB-2001; 2001US-269308P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.

XX  
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 XX

DR WPI; 2001-611495/70.  
 DR N-PSDB; AAS5437.

XX  
 PT New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -

XX  
 PS Example 3; Seq ID No 13171; 51lpp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 398 AA;

Query Match 86.2%; Score 106; DB 22; Length 398;  
 Best Local Similarity 85.7%; Pred. No. 6.7e-09;  
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 |||||:|||||:|||||  
 Db 3 kekydrskphvntgtnvdh 23

RESULT 15

AAR20243  
 ID AAR20243 standard; Protein; 396 AA.

XX  
 AC AAR20243;

DT 15-APR-1992 (first entry)

DE Sequence of translation elongation factor Tu2 encoded by tuf2 gene.

XX Elfamycin resistant actinomycetes; antibiotic resistant;  
 KW elongation factor.

XX Streptomyces ramocissimus.

OS EP466251-A.

PN 15-JAN-1992.

XX 02-JUL-1991; 91EP-0201702.

PR 02-JUL-1991; 91EP-0201702.









PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 82.9%; Score 102; DB 21; Length 465;  
Best Local Similarity 94.7%; Pred. No. 3.5e-08;  
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 KFNRTKPHVNIQTIGHVDH 22  
||| ||||| ||||| ||||| |||||  
Db 61 kferckphvniqtighvdh 79

## RESULT 23

AAU33391  
ID AAU33391 standard; Protein; 395 AA.

XX AC AAU33391;

XX DT 14-FEB-2002 (first entry)

XX DE Enterococcus faecalis cellular proliferation protein #27.

XX KW Antisense; prokaryotic cellular proliferation protein;

XX KW antibiotic; antibacterial; drug design.

XX OS Enterococcus faecalis.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207727P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

XX DR N-PSDB; AAS51250.

XX PT New polynucleotides for the identification and development of

XX PT antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 4887; 511pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to

XX CC prokaryotic cellular proliferation, their use in identifying the

XX CC genes, their use in the discovery of novel antibiotics, the essential

XX CC genes themselves and the encoded proteins. The prokaryotes used are

XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

XX CC invention is also useful for the identification of potential new targets

XX CC for antibiotic development. The antisense nucleic acids can also be used

XX CC to identify proteins used in proliferation, to express these proteins,

XX CC and to obtain antibodies capable of binding to the expressed proteins.

XX CC The proteins can be used to screen compounds in rational drug discovery

XX CC programmes. The antisense nucleic acid sequence is also useful to screen

XX CC for homologous nucleic acids which are required for cell proliferation in

XX CC a wide variety of organisms. The present sequence represents an

XX CC essential prokaryotic cellular proliferation protein.

XX CC Note: The sequence data for this patent did not form part

XX CC of the printed specification, but was obtained in electronic

XX CC format directly from WIPO at

XX CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 395 AA;

Query Match 82.1%; Score 101; DB 22; Length 395;  
Best Local Similarity 85.7%; Pred. No. 4.3e-08;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 KEKFNRTKPHVNIQTIGHVDH 22  
||| ||| ||| ||| ||| ||| |||

Db 3 kekfnrtkphvniqtighvdh 23

## RESULT 24

AAU34889  
ID AAU34889 standard; Protein; 395 AA.

XX AC AAU34889;

XX DT 14-FEB-2002 (first entry)

XX DE Enterococcus faecalis cellular proliferation protein #176.

XX KW Antisense; prokaryotic cellular proliferation protein;

XX KW antibiotic; antibacterial; drug design.

XX OS Enterococcus faecalis.

XX PN WO200170955-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US09180.

XX PR 21-MAR-2000; 2000US-191078P.

XX PR 23-MAY-2000; 2000US-206848P.

XX PR 26-MAY-2000; 2000US-207727P.

XX PR 23-OCT-2000; 2000US-242578P.

XX PR 27-NOV-2000; 2000US-253625P.

XX PR 22-DEC-2000; 2000US-257931P.

XX PR 16-FEB-2001; 2001US-269308P.

XX PA (ELIT-) ELITRA PHARM INC.

XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

XX DR N-PSDB; AAS52748.

XX PT New polynucleotides for the identification and development of

XX PT antibiotics, comprise sequences of antisense nucleic acids -

XX PS Example 3; Seq ID No 10482; 511pp; English.

XX CC The invention relates to antisense inhibitors of genes essential to

XX CC prokaryotic cellular proliferation, their use in identifying the

XX CC genes, their use in the discovery of novel antibiotics, the essential

XX CC genes themselves and the encoded proteins. The prokaryotes used are

XX CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella

XX CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

XX CC invention is also useful for the identification of potential new targets

XX CC for antibiotic development. The antisense nucleic acids can also be used

XX CC to identify proteins used in proliferation, to express these proteins,

XX CC and to obtain antibodies capable of binding to the expressed proteins.

XX CC The proteins can be used to screen compounds in rational drug discovery

XX CC programmes. The antisense nucleic acid sequence is also useful to screen

XX CC for homologous nucleic acids which are required for cell proliferation in

XX CC a wide variety of organisms. The present sequence represents an

XX CC essential prokaryotic cellular proliferation protein.

XX CC Note: The sequence data for this patent did not form part

XX CC of the printed specification, but was obtained in electronic

XX CC format directly from WIPO at

XX CC ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 395 AA;

Query Match 82.1%; Score 101; DB 22; Length 395;  
Best Local Similarity 85.7%; Pred. No. 4.3e-08;  
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
||||:|:| ||||| |||||  
Db 3 kekfdrrskshvniptighvdh 23

RESULT 25

AAU66514  
ID AAU66514 standard; Protein; 132 AA.

XX AC AAU66514;

XX DT 27-FEB-2002 (first entry)

XX DE Propionibacterium acnes immunogenic protein #27410.

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.

XX OS Propionibacterium acnes.

XX PN WO200101581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

XX PA (CORI-) CORIXA CORP.

XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX DR WPI; 2001-616774/71.

XX DR N-PSDB; AAS59738.

XX PS Example 1; SEQ ID No 27709; 1069pp; English.

CC Sequences AAU39105-AAU58017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence, for example, by enzyme linked immunosorbent assay (ELISA).

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 132 AA;

Query Match 81.3%; Score 100; DB 22; Length 132;  
Best Local Similarity 81.0%; Pred. No. 1.8e-08;  
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
||| ||||| ||||| |||||  
Db 42 kakferrkphcniptighdh 62

RESULT 26

AAU34672  
ID AAU34672 standard; Protein; 394 AA.

XX AC AAU34672;

XX DT 13-SEP-1999 (first entry)

XX DE Chlamydia pneumoniae transmembrane protein sequence.

XX KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
KW vaccine; neutralising epitope.

XX OS Chlamydia pneumoniae.

XX PN WO9927105-A2.

XX PD 03-JUN-1999.

XX PF 20-NOV-1998; 98WO-IB01890.

XX PR 04-NOV-1998; 98US-0107078.

XX PR 21-NOV-1997; 97FR-0014673.

XX PA (GEST ) GENSET.

XX PI Griffais R;

XX DR WPI; 1999-357842/30.

XX PT Genome sequence of Chlamydia pneumoniae

XX PS Page 675; Disclosure; 1912pp; English.

CC AAU34584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see AAX91990) of Chlamydia pneumoniae. C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart disease, sarcoidosis, sinusitis, purulent otitis media, erythema nodosum or pharyngitis. The polypeptides encoded by the open reading frames of the C. pneumoniae genome (see AAU34584-Y35879) can be used in immunogenic compositions as vaccines. Vectors containing C. pneumoniae nucleotide sequences can also be used as immunogenic compositions, CC especially where the vector directs the expression of a neutralising CC epitope of C. pneumoniae.

XX SQ Sequence 394 AA;

Query Match 81.3%; Score 100; DB 20; Length 394;  
Best Local Similarity 81.0%; Pred. No. 6.2e-08;  
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
||| ||||| ||||| |||||  
Db 3 kekfnrktphcniptighdh 23

RESULT 27

AAU38889  
ID AAU38889 standard; Protein; 394 AA.  
XX  
AC AAU38889;  
XX  
DT 16-JAN-2002 (first entry)  
XX  
DE C. pneumoniae CT322 homologue CPn0074.  
XX  
KW Chlamydia; sexually transmitted disease; PID; antibacterial;  
KW pelvic inflammatory disease; antigen; trachoma; gynecological;  
KW acute respiratory tract infection; atherosclerosis; male infertility;  
KW coronary heart disease.  
XX  
OS Chlamydia pneumoniae.  
XX  
PN WO200181379-A2.  
XX  
XX 01-NOV-2001.  
XX  
PF 23-APR-2001; 2001WO-US13081.  
XX  
PR 21-APR-2000; 2000US-198853P.  
PR 20-JUL-2000; 2000US-219752P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Bhatia A, Probst P, Stromberg EJ;  
XX  
DR WPI; 2001-616771/71.  
DR N-PSDB; AAS57020.  
XX  
XX New polynucleotide for treating Chlamydia infections encodes a  
PT polynucleotides containing an immunogenic portion of a Chlamydia  
PT antigen -  
XX  
XX Disclosure; Page 147; 208pp; English.  
XX  
CC The invention relates to isolated polynucleotide encoding at least  
CC a partial Chlamydia protein which is an antigenic fragment, or the  
CC complements, fragments, homologues and variants, and antibodies  
CC raised against the antigenic proteins (or fragments). The nucleic  
CC acids, proteins and antibodies are used to diagnose and treat Chlamydia  
CC infections (e.g. a sexually transmitted disease, pelvic inflammatory  
CC disease (PID), acute respiratory tract infection, trachoma,  
CC atherosclerosis and coronary heart disease) in a patient, and in  
CC the treatment of male infertility. The compounds of the invention are  
CC also useful for detecting the presence of Chlamydia in a patient, and  
CC stimulating and/or expanding T cells specific for a Chlamydia protein.  
CC The present sequence represents a Chlamydia antigen.  
XX  
SQ Sequence 394 AA;  
  
Query Match 81.3%; Score 100; DB 22; Length 394;  
Best Local Similarity 81.0%; Pred. No. 6.2e-08;  
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 2 KEKFNRTKPHVNIQTIGHVDH 22  
||| | |||:|||||  
Db 3 ketfqrnkphnigtighvdh 23  
  
RESULT 28  
AAU38909  
ID AAU38909 standard; Protein; 394 AA.  
XX  
AC AAU38909;  
XX  
DT 16-JAN-2002 (first entry)  
XX  
DE C. trachomatis CT322 protein.  
XX

KW Chlamydia; sexually transmitted disease; PID; antibacterial;  
KW pelvic inflammatory disease; antigen; trachoma; gynecological;  
KW acute respiratory tract infection; atherosclerosis; male infertility;  
KW coronary heart disease.  
XX  
OS Chlamydia trachomatis serovar D.  
XX  
PN WO200181379-A2.  
XX  
XX 01-NOV-2001.  
XX  
PF 23-APR-2001; 2001WO-US13081.  
XX  
PR 21-APR-2000; 2000US-198853P.  
PR 20-JUL-2000; 2000US-219752P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
PI Bhatia A, Probst P, Stromberg EJ;  
XX  
DR WPI; 2001-616771/71.  
DR N-PSDB; AAS57041.  
XX  
XX New polynucleotide for treating Chlamydia infections encodes a  
PT polynucleotides containing an immunogenic portion of a Chlamydia  
PT antigen -  
XX  
XX Claim 13; Page 182-183; 208pp; English.  
XX  
CC The invention relates to isolated polynucleotide encoding at least  
CC a partial Chlamydia protein which is an antigenic fragment, or the  
CC complements, fragments, homologues and variants, and antibodies  
CC raised against the antigenic proteins (or fragments). The nucleic  
CC acids, proteins and antibodies are used to diagnose and treat Chlamydia  
CC infections (e.g. a sexually transmitted disease, pelvic inflammatory  
CC disease (PID), acute respiratory tract infection, trachoma,  
CC atherosclerosis and coronary heart disease) in a patient, and in  
CC the treatment of male infertility. The compounds of the invention are  
CC also useful for detecting the presence of Chlamydia in a patient, and  
CC stimulating and/or expanding T cells specific for a Chlamydia protein.  
CC The present sequence represents a Chlamydia antigen.  
XX  
SQ Sequence 394 AA;  
  
Query Match 81.3%; Score 100; DB 22; Length 394;  
Best Local Similarity 81.0%; Pred. No. 6.2e-08;  
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 2 KEKFNRTKPHVNIQTIGHVDH 22  
||| | |||:|||||  
Db 3 ketfqrnkphnigtighvdh 23  
  
RESULT 29  
AAU36864  
ID AAU36864 standard; Protein; 397 AA.  
XX  
AC AAU36864;  
XX  
DT 07-OCT-1999 (first entry)  
XX  
DE Protein involved in transcription, translation and/or maturation.  
XX  
KW Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;  
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KW Bartholinitis; pneumopathy; venereal lymphogranulomatosis.  
XX  
OS Chlamydia trachomatis.  
XX  
XX WO9928475-A2.  
XX

```

PD 10-JUN-1999.
XX
XX 27-NOV-1998; 98WO-IB01939.
XX
XX 04-NOV-1998; 98US-0107077.
XX 28-NOV-1997; 97FR-0015041.
XX 17-DEC-1997; 97FR-0016034.
XX
XX (GEST ) GENSET.
XX
XX Griffais R;
PI
XX WPI; 1999-371125/31.
XX
XX Genome sequence of Chlamydia trachomatis
XX
XX Disclosure; Page 735-736; 1755pp; English.
XX
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome
XX of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as
XX vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
XX can also be used to control growth of the microorganism. Chlamydia
XX trachomatis is responsible for a large number of diseases, e.g. eye
XX diseases such as conventional trachoma, nongonococcal urethritis,
XX paratrachoma, and inclusion conjunctivitis; genital diseases such as
XX nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
XX perihpatitis, Bartholinitis; pneumopathy in breast feeding infants;
XX and venereal lymphogranulomatosis. The polypeptides of the invention
XX may be of use in treating these diseases.
XX
XX SQ Sequence 397 AA;

Query Match 81.3%; Score 100; DB 20; Length 397;
Best Local Similarity 81.0%; Pred. No. 6.2e-08;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
   ||| ||| ||| ||| ||| |||
Db 8 ketfrnkphnigtighvdh 28

RESULT 30
AAU39992
ID AAU39992 standard; Protein; 436 AA.
XX
XX AAU39992;
XX
XX 13-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #888.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX
XX WO200181581-A2.
XX
XX 01-NOV-2001.
XX
XX 20-APR-2001; 2001WO-US12865.
XX
XX 21-APR-2000; 2000US-199047P.
XX 02-JUN-2000; 2000US-208841P.
XX 07-JUL-2000; 2000US-216747P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

```

---

```

XX WPI; 2001-616774/71.
DR N-PSDB; AAS59510.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX vaccinating against and diagnosing infections, especially useful for
XX treating acne vulgaris -
XX
XX Claim 6; SEQ ID No 1187; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 436 AA;

Query Match 81.3%; Score 100; DB 22; Length 436;
Best Local Similarity 81.0%; Pred. No. 6.9e-08;
Matches 17; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
   ||| ||| ||| ||| ||| |||
Db 42 kakfertkphcnigtighdh 62

RESULT 31
AAG82109
ID AAG82109 standard; Protein; 310 AA.
XX
XX AAG82109;
XX
XX 03-SEP-2001 (first entry)
XX
XX S. epidermidis open reading frame protein sequence SEQ ID NO:1312.
XX
XX Staphylococcus epidermidis SR1 strain; infection; diagnosis;
XX vaccination; endocarditis.
XX
XX Staphylococcus epidermidis.
XX
XX WO200134809-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000WO-US30782.
XX
XX 09-NOV-1999; 99US-0164258.
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Kimmerly WJ;
XX
XX WPI; 2001-316495/33.
DR N-PSDB; AAH52959.
XX

```

PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*,  
 XX useful for vaccinating against infections, e.g. endocarditis -  
 PS Claim 18; Page 374; 218pp; English.  
 CC AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*.  
 CC (I) and (II) can have antibacterial activity and therefore can be used  
 CC in vaccination. The nucleic acids (I) may be used to produce the  
 CC *S. epidermidis* polypeptides (II) via the production of vectors  
 CC containing them which are used to produce hosts cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be  
 CC used to vaccinate subjects and to raise antibodies against the bacteria.  
 CC The polypeptides may also be used to assay for other inhibitors of their  
 CC activity and therefore identify compounds that may be used for the  
 CC treatment of *S. epidermidis* infections, e.g. endocarditis. AAH53971 to  
 CC AAH55090 represent specifically claimed *S. epidermidis* genomic DNA  
 CC polynucleotide sequences from the present invention. AAH55091 to  
 CC AAH55098 represent oligonucleotide sequences and primers which are used  
 CC in the exemplification of the present invention.  
 CC N.B. The present invention specifically claims all the polynucleotide  
 CC sequences given in the sequence listing of the present specification,  
 CC however the sequence listing only goes up to SEQ ID NO:4454 so even  
 CC though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,  
 CC no sequences are present for SEQ ID NO:4455 to 4464.

XX Sequence 310 AA;

Query Match 78.9%; Score 97; DB 22; Length 310;  
 Best Local Similarity 81.0%; Pred. NO. 1.4e-07;  
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||:|:| | |||||  
 Db 3 kekfdsrkehanigtighvdh 23

RESULT 32  
 AAU96769

ID AAU96769 standard; Protein; 394 AA.

XX AAU96769;

DT 26-SEP-2000 (first entry)

DE *Staphylococcus aureus* elongation factor Tu.

XX EF-Tu; elongation factor; antibiotic; anti-microbial.

XX *Staphylococcus aureus*.

XX WO200037495-A1.

PD 29-JUN-2000.

PF 08-DEC-1999; 99WO-US29098.

PR 21-DEC-1998; 98US-0218197.

FA (SMIK ) SMITHKLINE BEECHAM CORP.

XX Kallender H, Van Horn S;

XX WPI; 2000-442640/38.

DR N-PSDB; AA551329.

XX Novel elongation factors of *Staphylococcus aureus* useful as antibiotics  
 PT for treating microbial diseases and infections, comprise a specific  
 PT amino acid sequence

XX Claim 1; Page 36-37; 40pp; English.

XX

CC This is the *Staphylococcus aureus* elongation factor Tu (EF-Tu). EF-Tu is  
 CC useful for treating an individual in need of enhanced activity or  
 CC expression of EF-Tu. EF-Tu is useful as an antibiotic for treating  
 CC microbial infections and diseases. An antagonist to EF-Tu is useful for  
 CC treating an individual in need of inhibiting the activity or expression  
 CC of EF-Tu.

XX Sequence 394 AA;

Query Match 78.9%; Score 97; DB 21; Length 394;  
 Best Local Similarity 81.0%; Pred. NO. 1.9e-07;  
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||:|:| | |||||  
 Db 3 kekfdsrkehanigtighvdh 23

RESULT 33  
 AAU34078

ID AAU34078 standard; Protein; 394 AA.

XX AAU34078;

DT 14-FEB-2002 (first entry)

DE *Staphylococcus aureus* cellular proliferation protein #354.

XX Antisense: prokaryotic cellular proliferation protein;

XX antibiotic; antibacterial; drug design.

OS *Staphylococcus aureus*.

XX WO200170955-A2.

XX 27-SEP-2001.

PF 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-208648P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

XX Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

DR N-PSDB; AA551937.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID NO 5574; 51pp; English.

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella*  
 CC *pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery  
 CC programmes. The antisense nucleic acid sequence is also useful to screen

CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 394 AA;

Query Match 78.9%; Score 97; DB 22; Length 394;  
 Best Local Similarity 81.0%; Pred. No. 1.9e-07;  
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||:|:| | |||||  
 Db 3 kekfrskehanigtighvdh 23

RESULT 34  
 AAU36849  
 ID AAU36849 standard; Protein; 394 AA.  
 AC  
 AC AAU36849;  
 XX  
 XX 14-FEB-2002 (first entry)  
 DT  
 DT Staphylococcus aureus cellular proliferation protein #1019.  
 XX

XX Antisense; prokaryotic cellular proliferation protein;  
 KW antibiotic; antibacterial; drug design.  
 KW  
 XX Staphylococcus aureus.  
 OS  
 XX WO200170955-A2.  
 PN  
 XX 27-SEP-2001.  
 PD  
 XX  
 XX 21-MAR-2001; 2001WO-US09180.  
 PF  
 XX  
 XX 21-MAR-2000; 2000US-191078P.  
 PR  
 XX 23-MAY-2000; 2000US-206848P.  
 PR  
 XX 26-MAY-2000; 2000US-207727P.  
 PR  
 XX 23-OCT-2000; 2000US-242578P.  
 PR  
 XX 27-NOV-2000; 2000US-253625P.  
 PR  
 XX 22-DEC-2000; 2000US-257931P.  
 PR  
 XX 16-FEB-2001; 2001US-269308P.  
 PR

XX (ELIT-) ELITRA PHARM INC.  
 PA  
 PA Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;  
 PI Yamamoto RT, Xu HH;  
 PI  
 XX WPI; 2001-611495/70.  
 DR N-PSDB; AAS54708.  
 DR

XX New polynucleotides for the identification and development of  
 PT antibiotics, comprise sequences of antisense nucleic acids -  
 PT  
 XX Example 3; Seq ID No 12442; 51pp; English.  
 PS

XX The invention relates to antisense inhibitors of genes essential to  
 CC prokaryotic cellular proliferation, their use in identifying the  
 CC genes, their use in the discovery of novel antibiotics, the essential  
 CC genes themselves and the encoded proteins. The prokaryotes used are  
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella  
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The  
 CC invention is also useful for the identification of potential new targets  
 CC for antibiotic development. The antisense nucleic acids can also be used  
 CC to identify proteins used in proliferation, to express these proteins,  
 CC and to obtain antibodies capable of binding to the expressed proteins.  
 CC The proteins can be used to screen compounds in rational drug discovery

CC programmes. The antisense nucleic acid sequence is also useful to screen  
 CC for homologous nucleic acids which are required for cell proliferation in  
 CC a wide variety of organisms. The present sequence represents an  
 CC essential prokaryotic cellular proliferation protein.  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 XX Sequence 394 AA;

Query Match 78.9%; Score 97; DB 22; Length 394;  
 Best Local Similarity 81.0%; Pred. No. 1.9e-07;  
 Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22  
 ||||:|:| | |||||  
 Db 3 kekfrskehanigtighvdh 23

RESULT 35  
 AAG12253  
 ID AAG12253 standard; Protein; 77 AA.  
 XX  
 XX AAG12253;  
 AC  
 AC 17-OCT-2000 (first entry)  
 DT  
 DT Zea mays protein fragment SEQ ID NO: 11293.  
 XX

XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 KW  
 XX Zea mays subsp. mays.  
 OS  
 XX EP1033405-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825.  
 PR  
 XX 05-MAR-1999; 99US-0123180.  
 PR  
 XX 09-MAR-1999; 99US-0123548.  
 PR  
 XX 23-MAR-1999; 99US-0125788.  
 PR  
 XX 25-MAR-1999; 99US-0126264.  
 PR  
 XX 29-MAR-1999; 99US-0126785.  
 PR  
 XX 01-APR-1999; 99US-0127462.  
 PR  
 XX 06-APR-1999; 99US-0128234.  
 PR  
 XX 08-APR-1999; 99US-0128714.  
 PR  
 XX 16-APR-1999; 99US-0129845.  
 PR  
 XX 19-APR-1999; 99US-0130077.  
 PR  
 XX 21-APR-1999; 99US-0130449.  
 PR  
 XX 23-APR-1999; 99US-0130510.  
 PR  
 XX 28-APR-1999; 99US-0130891.  
 PR  
 XX 30-APR-1999; 99US-0131449.  
 PR  
 XX 30-APR-1999; 99US-0132048.  
 PR  
 XX 04-MAY-1999; 99US-0132407.  
 PR  
 XX 05-MAY-1999; 99US-0132484.  
 PR  
 XX 06-MAY-1999; 99US-0132485.  
 PR  
 XX 06-MAY-1999; 99US-0132486.  
 PR  
 XX 07-MAY-1999; 99US-0132487.  
 PR  
 XX 07-MAY-1999; 99US-0132863.  
 PR  
 XX 11-MAY-1999; 99US-0134256.  
 PR  
 XX 14-MAY-1999; 99US-0134218.  
 PR  
 XX 14-MAY-1999; 99US-0134219.  
 PR  
 XX 14-MAY-1999; 99US-0134221.  
 PR  
 XX 14-MAY-1999; 99US-0134370.  
 PR  
 XX 18-MAY-1999; 99US-0134768.  
 PR  
 XX 19-MAY-1999; 99US-0134941.  
 PR  
 XX 20-MAY-1999; 99US-0135124.



PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136382.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139494.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.



PR 06-AUG-1999;	99US-0147303.	Matches	16;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
PR 06-AUG-1999;	99US-0147416.	Qy	5	FNRTRPHVNICTIGHVDH	22						
PR 09-AUG-1999;	99US-0147493.				:						
PR 09-AUG-1999;	99US-0147935.	Db	13	frtkphvngtighvdh	30						
PR 10-AUG-1999;	99US-0148171.										
PR 11-AUG-1999;	99US-0148319.										
PR 12-AUG-1999;	99US-0148341.										
PR 13-AUG-1999;	99US-0148565.										
PR 13-AUG-1999;	99US-0148684.										
PR 16-AUG-1999;	99US-0149368.										
PR 17-AUG-1999;	99US-0149175.										
PR 18-AUG-1999;	99US-0149426.										
PR 20-AUG-1999;	99US-0149722.										
PR 20-AUG-1999;	99US-0149723.										
PR 20-AUG-1999;	99US-0149929.										
PR 23-AUG-1999;	99US-0149902.										
PR 23-AUG-1999;	99US-0149930.										
PR 25-AUG-1999;	99US-0150566.										
PR 26-AUG-1999;	99US-0150884.										
PR 27-AUG-1999;	99US-0151065.										
PR 27-AUG-1999;	99US-0151066.										
PR 27-AUG-1999;	99US-0151080.										
PR 30-AUG-1999;	99US-0151303.										
PR 31-AUG-1999;	99US-0151438.										
PR 01-SEP-1999;	99US-0151930.										
PR 07-SEP-1999;	99US-0152363.										
PR 10-SEP-1999;	99US-0153070.										
PR 13-SEP-1999;	99US-0153758.										
PR 15-SEP-1999;	99US-0154018.										
PR 16-SEP-1999;	99US-0154039.										
PR 20-SEP-1999;	99US-0154779.										
PR 22-SEP-1999;	99US-0155139.										
PR 23-SEP-1999;	99US-0155486.										
PR 24-SEP-1999;	99US-0155659.										
PR 28-SEP-1999;	99US-0156458.										
PR 29-SEP-1999;	99US-0156596.										
PR 04-OCT-1999;	99US-0157117.										
PR 05-OCT-1999;	99US-0157753.										
PR 06-OCT-1999;	99US-0157865.										
PR 07-OCT-1999;	99US-0158029.										
PR 08-OCT-1999;	99US-0158232.										
PR 12-OCT-1999;	99US-0158369.										
PR 13-OCT-1999;	99US-0159293.										
PR 13-OCT-1999;	99US-0159294.										
PR 13-OCT-1999;	99US-0159295.										
PR 14-OCT-1999;	99US-0159329.										
PR 14-OCT-1999;	99US-0159330.										
PR 14-OCT-1999;	99US-0159331.										
PR 14-OCT-1999;	99US-0159637.										
PR 18-OCT-1999;	99US-0159638.										
PR 18-OCT-1999;	99US-0159584.										
PR 21-OCT-1999;	99US-0160741.										
PR 21-OCT-1999;	99US-0160767.										
PR 21-OCT-1999;	99US-0160768.										
PR 21-OCT-1999;	99US-0160770.										
PR 21-OCT-1999;	99US-0160814.										
PR 21-OCT-1999;	99US-0160815.										
PR 22-OCT-1999;	99US-0160980.										
PR 22-OCT-1999;	99US-0160981.										
PR 22-OCT-1999;	99US-0160981.										
PR 25-OCT-1999;	99US-0160989.										
PR 25-OCT-1999;	99US-0161404.										
PR 25-OCT-1999;	99US-0161405.										
PR 25-OCT-1999;	99US-0161406.										
PR 26-OCT-1999;	99US-0161359.										
PR 26-OCT-1999;	99US-0161360.										
PR 26-OCT-1999;	99US-0161361.										
PR 28-OCT-1999;	99US-0161920.										
PR 28-OCT-1999;	99US-0161922.										
PR 28-OCT-1999;	99US-0161993.										
PR 29-OCT-1999;	99US-0162142.										

PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 78.0%; Score 96; DB 21; Length 111;  
Best Local Similarity 88.9%; Pred.No. 6.8e-08;

Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 FNRTRPHVNIQTIGHVDH 22

Db 46 ftrtkphvnnvgtighvdh 63

RESULT 38  
AAG12474  
ID AAG12474 standard; Protein; 129 AA.  
XX AC AAG12474;  
XX DT 17-OCT-2000 (first entry)  
XX DE Zea mays protein fragment SEQ ID NO: 11601.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence; corn.  
XX OS Zea mays subsp. mays.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139118.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.

PR	20-AUG-1999;	99US-0149929.	XX	Zea mays protein fragment SEQ ID NO: 55934.
PR	23-AUG-1999;	99US-0149902.	DE	
PR	23-AUG-1999;	99US-0149930.	XX	
PR	25-AUG-1999;	99US-0150566.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	26-AUG-1999;	99US-0150884.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	27-AUG-1999;	99US-0151085.	KW	termination sequence; corn.
PR	27-AUG-1999;	99US-0151086.	XX	
PR	27-AUG-1999;	99US-0151080.	OS	Zea mays subsp. mays.
PR	30-AUG-1999;	99US-0151303.	XX	
PR	31-AUG-1999;	99US-0151438.	PN	EPI033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	06-SEP-2000.
PR	10-SEP-1999;	99US-0153070.	XX	
PR	13-SEP-1999;	99US-0153758.	PF	25-FEB-2000; 2000EP-0301439.
PR	15-SEP-1999;	99US-0154018.	XX	
PR	16-SEP-1999;	99US-0154039.	XX	25-FEB-1999; 99US-0121825.
PR	20-SEP-1999;	99US-0154779.	PR	05-MAR-1999; 99US-0123180.
PR	22-SEP-1999;	99US-0155119.	PR	09-MAR-1999; 99US-0123548.
PR	23-SEP-1999;	99US-0155486.	PR	23-MAR-1999; 99US-0125788.
PR	24-SEP-1999;	99US-0155659.	PR	25-MAR-1999; 99US-0126264.
PR	28-SEP-1999;	99US-0156458.	PR	29-MAR-1999; 99US-0126785.
PR	29-SEP-1999;	99US-0156596.	PR	01-APR-1999; 99US-0127462.
PR	04-OCT-1999;	99US-0157117.	PR	06-APR-1999; 99US-0128234.
PR	05-OCT-1999;	99US-0157753.	PR	08-APR-1999; 99US-0128714.
PR	06-OCT-1999;	99US-0157865.	PR	16-APR-1999; 99US-0129845.
PR	07-OCT-1999;	99US-0158029.	PR	19-APR-1999; 99US-0130077.
PR	08-OCT-1999;	99US-0158232.	PR	21-APR-1999; 99US-0130449.
PR	12-OCT-1999;	99US-0158369.	PR	23-APR-1999; 99US-0130510.
PR	13-OCT-1999;	99US-0159293.	PR	23-APR-1999; 99US-0130891.
PR	13-OCT-1999;	99US-0159294.	PR	28-APR-1999; 99US-0131449.
PR	13-OCT-1999;	99US-0159295.	PR	30-APR-1999; 99US-0132048.
PR	14-OCT-1999;	99US-0159329.	PR	30-APR-1999; 99US-0132407.
PR	14-OCT-1999;	99US-0159331.	PR	04-MAY-1999; 99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	05-MAY-1999; 99US-0132485.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999; 99US-0132486.
PR	18-OCT-1999;	99US-0159584.	PR	06-MAY-1999; 99US-0132487.
PR	21-OCT-1999;	99US-0160741.	PR	07-MAY-1999; 99US-0132863.
PR	21-OCT-1999;	99US-0160767.	PR	11-MAY-1999; 99US-0134256.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999; 99US-0134218.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999; 99US-0134219.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999; 99US-0134221.
PR	21-OCT-1999;	99US-0160815.	PR	14-MAY-1999; 99US-0134370.
PR	22-OCT-1999;	99US-0160980.	PR	18-MAY-1999; 99US-0134768.
PR	22-OCT-1999;	99US-0160981.	PR	19-MAY-1999; 99US-0134941.
PR	22-OCT-1999;	99US-0160989.	PR	20-MAY-1999; 99US-0135124.
PR	25-OCT-1999;	99US-0161404.	PR	21-MAY-1999; 99US-0135353.
PR	25-OCT-1999;	99US-0161405.	PR	24-MAY-1999; 99US-0135629.
PR	25-OCT-1999;	99US-0161406.	PR	25-MAY-1999; 99US-0136021.
PR	26-OCT-1999;	99US-0161359.	PR	27-MAY-1999; 99US-0136392.
PR	26-OCT-1999;	99US-0161360.	PR	28-MAY-1999; 99US-0136782.
PR	26-OCT-1999;	99US-0161361.	PR	01-JUN-1999; 99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	03-JUN-1999; 99US-0137528.
PR	28-OCT-1999;	99US-0161920.	PR	04-JUN-1999; 99US-0137502.
PR	28-OCT-1999;	99US-0161992.	PR	07-JUN-1999; 99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	08-JUN-1999; 99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999; 99US-0138540.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999; 99US-0138847.
PR	29-OCT-1999;	99US-0162142.	PR	10-JUN-1999; 99US-0139119.
PR	29-OCT-1999;	99US-0162142.	PR	14-JUN-1999; 99US-0139452.
PR	29-OCT-1999;	99US-0162142.	PR	16-JUN-1999; 99US-0139453.
PR	29-OCT-1999;	99US-0162142.	PR	17-JUN-1999; 99US-0139492.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139454.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139455.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139456.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139457.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139458.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139459.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139460.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139461.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139462.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139463.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139750.
PR	29-OCT-1999;	99US-0162142.	PR	18-JUN-1999; 99US-0139763.

Query Match 78.08; Score 96; DB 21; Length 129;  
Best Local Similarity 88.98; Pred. No. 8e-08;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 5 FNRKTPHVNICTIGHVDH 22  
Db 64 ftrtkphvnvgtlgnvdh 81  
RESULT 39  
AAG44634  
ID AAG44634 standard; Protein; 156 AA.  
XX  
AC AAG44634;  
XX  
DT 18-OCT-2000 (first entry)







PR	15-SEP-1999;	99US-01540318;
PR	16-SEP-1999;	99US-01540339;
PR	20-SEP-1999;	99US-01547799;
PR	22-SEP-1999;	99US-01551339;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-015659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;

Query Match 78.0%; Score 96; DB 21; Length 175;  
Best Local Similarity 88.9%; Pred. No. 1.1e-07;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      5  FNRTKPHVNI GTIGHVDH 22
          | | | | | | | | | | | | | |
Db     79  ftrtkphvnnvgtighvdh 96
```

Search completed: August 22, 2002, 07:44:04  
Job time: 63 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 22, 2002, 07:43:01 ; Search time 13.53 Seconds  
(without alignments)  
62.959 Million cell updates/sec

Title: US-09-488-737-1

Perfect score: 123

Sequence: 1 MKEKFNRTKPHVNIQTIGHVDH 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	95.9	399	1 EFTU_HELPY	P56003 helicobacte
2	112	91.1	394	1 EFTU_BACFR	P33165 bacteroides
3	111	90.2	393	1 EFTU_ECOLI	P02990 escherichia
4	111	90.2	393	1 EFTU_HAEIN	P43926 haemophilus
5	111	90.2	393	1 EFTU_SALTY	P21694 salmonella
6	111	90.2	394	1 EFTU_PASMU	P57939 pasteurella
7	111	90.2	394	1 EFTB_PASMU	P57966 pasteurella
8	109	88.6	409	1 EFTU_ODOSI	P49462 odontella s
9	109	88.6	410	1 EFTU_NEPOL	Q9Lkz5 nephroselmi
10	108	87.8	399	1 EFTU_CAMJE	O69303 campylobact
11	107	87.0	394	1 EFTU_NEIGO	P48864 neisseria g
12	107	87.0	394	1 EFTU_RECAM	O21245 reclinomona
13	107	87.0	396	1 EFTU_STIAG	P42479 stigmatella
14	107	87.0	400	1 EFTU_THEMA	P13537 thermotoga
15	106	86.2	394	1 EFTU_RICPR	P48865 rickettsia
16	106	86.2	394	1 EFTU_UREPA	P50068 ureaplasma
17	106	86.2	395	1 EFTU_BACST	O50306 bacillus st
18	106	86.2	396	1 EFTU_CORGL	P42439 corynebacte
19	106	86.2	396	1 EFTU_MYCLE	P30768 mycobacteri
20	106	86.2	396	1 EFTU_MYCTU	P13501 mycobacteri
21	106	86.2	396	1 EFTU_THICU	P42481 thiobacilli
22	106	86.2	397	1 EFTU_PSEAE	P09591 pseudomonas
23	106	86.2	397	1 EFTU_STRCJ	P95724 streptomyce
24	106	86.2	398	1 EFTU_STROR	P33170 streptococ
25	106	86.2	398	1 EFTU_STRPY	P82559 streptococ
26	105	85.4	391	1 EFTU_RHILO	Q981f7 rhizobium 1
27	105	85.4	394	1 EFTU_MYCGA	P18906 mycoplasma
28	105	85.4	397	1 EFT1_STRCO	P40174 streptomyce
29	105	85.4	397	1 EFT1_STRCU	Q53871 streptomyce
30	105	85.4	397	1 EFT1_STRRA	P29542 streptomyce
31	105	85.4	397	1 EFT2_STRRA	P29543 streptomyce
32	105	85.4	397	1 EFTU_STRAU	O33594 streptomyce
33	105	85.4	409	1 EFTU_EUGGR	P02991 euglena gra

## ALIGNMENTS

RESULT 1

```

EFTU_HELPY
ID EFTU_HELPY STANDARD; PRT; 399 AA.
AC P56003;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR TUF1 OR HP1205.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., B.A.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty S.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
RC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000626; AAC08250.1; -
DR HSP; P02990; 1ETU.
DR TIGR; HP1205; -.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 399 AA; 43648 MW; 502751637BDA2707 CRC64;

Query Match 95.98; Score 118; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 6.1e-12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KEKENRTKPHVNIQTIGHVDH 22
Db 3 KEKENRTKPHVNIQTIGHVDH 23

RESULT 2

```

```

EFTU_BACFR
ID EFTU_BACFR STANDARD; PRT; 394 AA.
AC P33165;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Bacteroides fragilis.
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
OC Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90240875; PubMed=2110445;
RA Ludwig W., Weizenegger M., Betzel D., Leidel E., Lenz T., Ludvigsen A.,
RA Moellenhoff D., Wenzig P., Schliefer K.H.;
RT "Complete nucleotide sequences of seven eubacterial genes coding for
RT the elongation factor Tu: functional, structural and phylogenetic
RT evaluations."
RL Arch. Microbiol. 153:241-247(1990).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC PIR; B06663; B06663.
DR HSP; P02990; 1ETU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFACOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 394 AA; 43580 MW; 7B4C6FD208323149 CRC64;

Query Match 91.18; Score 112; DB 1; Length 394;
Best Local Similarity 95.28; Pred. No. 5.9e-11;
Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KEKENRTKPHVNIQTIGHVDH 22
Db 3 KEKENRTKPHVNIQTIGHVDH 23

RESULT 3
EFTU_ECOLI
ID EFTU_ECOLI STANDARD; PRT; 393 AA.
AC P02990;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu) (P-43).
GN (TUF1 OR B3339) AND (TUF2 OR B3980).
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A. (TUF1).
RX MEDLINE=8116558; PubMed=7011904;
RA An G., Friesen J.D.;
RT "The nucleotide sequence of tufB and four nearby trna structural
RT genes of Escherichia coli."

```

RL Gene 12:33-39(1980).  
RN [2]  
RP SEQUENCE FROM N.A. (TUFb).  
RC STRAIN-K12 / MG1655;  
RX MEDLINE-94089392; PubMed-8265357;  
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,  
RA Daniels D.L.;  
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the  
RT region from 89,2 to 92,8 minutes";  
RL Nucleic Acids Res. 21:5408-5417(1993).  
RN [3]  
RP SEQUENCE FROM N.A. (TUFa).  
RX MEDLINE-81165557; PubMed-7011903;  
RA Yokota T., Sugisaki H., Takanami M., Kaziro Y.;  
RT "The nucleotide sequence of the cloned tufA gene of Escherichia  
RT coli";  
RL Gene 12:25-31(1980).  
RN [4]  
RP SEQUENCE FROM N.A. (TUFa).  
RC STRAIN-K12 / MG1655;  
RX MEDLINE-97426617; PubMed-9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12";  
RL Science 277:1453-1474(1997).  
RN [5]  
RP SEQUENCE.  
RX MEDLINE-81003875; PubMed-6997043;  
RA Jones M.D., Petersen T.E., Nielsen K.M., Magnusson S.,  
RA Sottrup-Jensen L., Gausing K., Clark B.F.C.;  
RT "The complete amino-acid sequence of elongation factor Tu from  
RT Escherichia coli";  
RL Eur. J. Biochem. 108:507-526(1980).  
RN [6]  
RP SEQUENCE.  
RX MEDLINE-81264196; PubMed-7021545;  
RA Laursen R.A., L'Italien J.J., Nagarkatti S., Miller D.L.;  
RT "The amino acid sequence of elongation factor Tu of Escherichia coli.  
RT The complete sequence";  
RL J. Biol. Chem. 256:8102-8109(1981).  
RN [7]  
RP SEQUENCE OF 1-20 FROM N.A. (TUFb).  
RX MEDLINE-82124558; PubMed-7035813;  
RA Miyajima A., Shibuya M., Kuchino Y., Kaziro Y.;  
RT "Transcription of the E. coli tufB gene: cotranscription with four  
RT tRNA genes and inhibition by  
RT guanosine-5'-diphosphate-3'-diphosphate";  
RL Mol. Gen. Genet. 183:13-19(1981).  
RN [8]  
RP SEQUENCE OF 1-12 AND 310-321.  
RC STRAIN-K12 / EMG2;  
RX MEDLINE-97443975; PubMed-9298646;  
RA Link A.J., Robison K., Church G.M.;  
RT "Comparing the predicted and observed properties of proteins encoded  
RT in the genome of Escherichia coli K-12";  
RL Electrophoresis 18:1259-1313(1997).  
RN [9]  
RP PHOSPHORYLATION SITE THR-382, AND SEQUENCE OF 289-303 AND 382-390.  
RX MEDLINE-93107067; PubMed-8416965;  
RA Lippmann C., Lindschau C., Vijgenboom E., Schroeder W., Bosch L.,  
RA Erdmann V.A.;  
RT "Prokaryotic elongation factor Tu is phosphorylated in vivo";  
RL J. Biol. Chem. 268:601-607(1993).  
RN [10]  
RP PARTIAL SEQUENCE OF 152-175 AND 261-289.  
RC STRAIN-B/R;  
RX MEDLINE-91216980; PubMed-2022614;  
RA Young C.C., Bernlohr R.W.;  
RT "Elongation factor Tu is methylated in response to nutrient  
RT deprivation in Escherichia coli";  
RL J. Bacteriol. 173:3096-3100(1991).

RN [11]  
RP METHYLATION.  
RX MEDLINE-80069392; PubMed-389663;  
RA L'Italien J.J., Laursen R.A.;  
RT "Location of the site of methylation in elongation factor Tu.";  
RL FEBS Lett. 107:359-362(1979).  
RN [12]  
RP MUTAGENESIS OF VAL-20.  
RX MEDLINE-90060119; PubMed-2684669;  
RA Jacquet E., Parmeggiani A.;  
RT "Substitution of Val20 by Gly in elongation factor Tu. Effects on the  
RT interaction with elongation factors Ts, aminoacyl-tRNA and  
RT ribosomes";  
RL Eur. J. Biochem. 185:341-346(1989).  
RN [13]  
RP SEQUENCE OF 75-89, AND MUTAGENESIS OF PRO-82.  
RX MEDLINE-90216700; PubMed-2157708;  
RA Cool R.H., Jensen M., Jonak J., Clark B.F.C., Parmeggiani A.;  
RT "Substitution of proline 82 by threonine induces autophosphorylating  
RT activity in GTP-binding domain of elongation factor Tu.";  
RL J. Biol. Chem. 265:6744-6749(1990).  
RN [14]  
RP MUTAGENESIS OF ASP-138.  
RX MEDLINE-88007508; PubMed-3308869;  
RA Hwang Y.-W., Miller D.L.;  
RT "A mutation that alters the nucleotide specificity of elongation  
RT factor Tu, a GTP regulatory protein";  
RL J. Biol. Chem. 262:13081-13085(1987).  
RN [15]  
RP MUTAGENESIS OF LYS-136.  
RX MEDLINE-89255270; PubMed-2498311;  
RA Hwang Y.-W., Sanchez A., Miller D.L.;  
RT "Mutagenesis of bacterial elongation factor Tu at lysine 136. A  
RT conserved amino acid in GTP regulatory proteins";  
RL J. Biol. Chem. 264:8304-8309(1989).  
RN [16]  
RP MUTAGENESIS.  
RX MEDLINE-90025117; PubMed-2508560;  
RA Hwang Y.-W., McCabe P.G., Innis M.A., Miller D.L.;  
RT "Site-directed mutagenesis of the GDP binding domain of bacterial  
RT elongation factor Tu.";  
RL Arch. Biochem. Biophys. 274:394-403(1989).  
RN [17]  
RP MUTAGENESIS OF HIS-19; GLN-114 AND GLU-348.  
RX MEDLINE-98136155; PubMed-9468511;  
RA Zhang Y., Yu N.J., Spremulli L.L.;  
RT "Mutational analysis of the roles of residues in Escherichia coli  
RT elongation factor Ts in the interaction with elongation factor Tu.";  
RL J. Biol. Chem. 273:4556-4562(1998).  
RN [18]  
RP CHARACTERIZATION OF MUTANT ASP-222.  
RX MEDLINE-97133305; PubMed-8978702;  
RA Vorstenbosch E., Pape T., Rodnina M.V., Kraal B., Wintermeyer W.;  
RT "The G222D mutation in elongation factor Tu inhibits the  
RT codon-induced conformational changes leading to GTPase activation on  
RT the ribosome";  
RL EMBO J. 15:6766-6774(1996).  
RN [19]  
RP VARIANTS RESISTANT TO KIRROMYCIN.  
RX MEDLINE-95045380; PubMed-7525272;  
RA Mesters J.R., Zeef L.A.H., Hilgenfeld R., de Graaf J.M., Kraal B.,  
RA Bosch L.;  
RT "The structural and functional basis for the kirromycin resistance of  
RT mutant EF-Tu species in Escherichia coli";  
RL EMBO J. 13:4877-4885(1994).  
RN [20]  
RP VARIANTS RESISTANT TO PULVOMYCIN.  
RX MEDLINE-95045403; PubMed-7957075;  
RA Zeef L.A.H., Bosch L., Anborgh P.H., Cetin R., Parmeggiani A.,  
RA Hilgenfeld R.;  
RT "Pulvomycin-resistant mutants of E.coli elongation factor Tu";  
RL EMBO J. 13:5113-5120(1994).  
RN [21]

```
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE=92177415; PubMed=1542116;
RA Kjeldgaard M., Nyborg J.;
RT "Refined structure of elongation factor EF-Tu from Escherichia coli.";
RL J. Mol. Biol. 223:721-742(1992).
RN [22]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF COMPLEX WITH EF-TS.
RX MEDLINE=96170031; PubMed=8596629;
RA Kawashima T., Berthet-Colominas C., Wulff M., Cusack S., Leberman R.;
RT "The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5-A
resolution.";
RL Nature 379:511-518(1996).
RN [23]
RP ERRATUM.
RA Kawashima T., Berthet-Colominas C., Wulff M., Cusack S., Leberman R.;
RL Nature 381:172-172(1996).
RN [24]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=97094972; PubMed=8939740;
RA Abel K., Toder M.D., Hilgenfeld R., Jurnak F.;
RT "An alpha to beta conformational switch in EF-Tu.";
RL Structure 4:1153-1159(1996).
RN [25]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=99141376; PubMed=9918724;
RA Song H., Parsons M.R., Rowse S., Leonard G., Phillips S.E.V.;
RT "Crystal structure of intact elongation factor EF-Tu from Escherichia
coli in GDP conformation at 2.05-A resolution.";
RL J. Mol. Biol. 285:1245-1256(1999).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- FUNCTION: MAY PLAY AN IMPORTANT REGULATORY ROLE IN CELL GROWTH AND
CC IN THE BACTERIAL RESPONSE TO NUTRIENT DEPRIVATION.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, AND MEMBRANE-ASSOCIATED.
CC -!- MISCELLANEOUS: THIS CHAIN IS ALSO USED IN BACTERIOPHAGE Q-BETA RNA
Query Match 90.2%; Score 111; DB 1; Length 393;
Best Local Similarity 90.5%; Pred. No. 8.6e-11;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 2 KEKFNRTKPHVNIPTIGHVDH 22
Db 2 KEKFNRTKPHVNVGTIGHVDH 22
|||||
RESULT 4
EFTU_HAEIN STANDARD; PRT; 393 AA.
AC P43926;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN (TUFA OR HI0578) AND (TUFB OR HI0632).
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
```

```
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U32739; AAC22236.1; -.
CC EMBL; U32746; AAC22292.1; -.
CC DR HSSP; P02990; 1ETU.
CC DR TIGR; HI0578; -.
CC DR TIGR; HI0632; -.
CC DR InterPro; IPR000795; GTP_EFTU.
CC DR InterPro; IPR004161; GTP_EFTU_D2.
CC DR InterPro; IPR004160; GTP_EFTU_D3.
CC DR Pfam; PF00009; GTP_EFTU; 1.
CC DR Pfam; PF03144; GTP_EFTU_D2; 1.
CC DR Pfam; PF03143; GTP_EFTU_D3; 1.
CC DR PRINTS; PR00315; ELONGATNFCT.
CC DR PROSITE; PS00301; EFACITOR_GTP; 1.
CC DR Elongation factor; Protein biosynthesis; GTP-binding;
CC Complete proteome.
CC INIT_MET 0 0 BY SIMILARITY.
CC NP_BIND 18 25 GTP (BY SIMILARITY).
CC FT NP_BIND 84 84 GTP (BY SIMILARITY).
CC FT NP_BIND 135 138 GTP (BY SIMILARITY).
CC FT BINDING 81 81 ASSOCIATED WITH AMINOACYL-TRNA BINDING
CC (BY SIMILARITY).
CC FT BINDING 137 137 ASSOCIATED WITH GUANOSINE NUCLEOTIDE
CC BINDING ACTIVITIES (BY SIMILARITY).
CC SEQUENCE 393 AA; 43223 MW; 2430107FD71A038C CRC64;
Query Match 90.2%; Score 111; DB 1; Length 393;
Best Local Similarity 90.5%; Pred. No. 8.6e-11;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 2 KEKFNRTKPHVNIPTIGHVDH 22
Db 2 KEKFNRTKPHVNVGTIGHVDH 22
|||||
RESULT 5
EFTU_SALTY STANDARD; PRT; 393 AA.
ID EFTU_SALTY
AC P21694;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN (TUFA OR STM3445 OR STY4353) AND (TUFB OR STM4146 OR STY3739).
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
SEQUENCE FROM N.A.
RC SPECIES=S. typhimurium; STRAIN=LT2;
RX MEDLINE=91002658; PubMed=2207156;
RA Tuohy T.M.F., Thompson S., Gesteland R.F., Hughes D., Atkins J.F.;
```

"The role of EF-Tu and other translation components in determining translocation step size.";  
 Biochim. Biophys. Acta 1050:274-278(1990).  
 [2]  
 SEQUENCE FROM N.A.  
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RA MEDLINE=21534948; PubMed=11677609;  
 RX McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 Waterston R., Wilson R.K.;  
 RA "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 LT2.";  
 RL Nature 413:852-856(2001).  
 [3]  
 SEQUENCE FROM N.A.  
 RC SPECIES=S.typhi; STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
 Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 Cronin A., Davis P., Davies R.W., Dowd L., White N., Farrar J.,  
 Feltwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,  
 Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 Whitehead S., Barrett B.G.;  
 RA "Complete genome sequence of a multiple drug resistant *Salmonella*  
 enterica serovar Typhi CT18.";  
 RL Nature 413:848-852(2001).  
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 BIOSYNTHESIS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X55116; CA38912.1; -;  
 DR EMBL; X55117; CA38913.1; -;  
 DR EMBL; AE008858; AAL22308.1; -;  
 DR EMBL; AE008893; AAL22974.1; -;  
 DR EMBL; AL627279; CAD09494.1; -;  
 DR EMBL; AL627281; CAD08168.1; -;  
 DR PIR; S13560; S13560.  
 DR PIR; S13561; S13561.  
 DR HSSP; P02990; LEFC.  
 DR StyGene; SG10403; tufa.  
 DR StyGene; SG10404; tufB.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFACFOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding; Phosphorylation;  
 KW Methylation; Complete proteome.  
 FT INIT\_MET 0 BY SIMILARITY.  
 FT NP\_BIND 18 25 GTP (BY SIMILARITY).  
 FT NP\_BIND 80 84 GTP (BY SIMILARITY).  
 FT NP\_BIND 135 138 GTP (BY SIMILARITY).  
 FT MOD\_RES 56 56 METHYLATION (BY SIMILARITY).  
 FT MOD\_RES 382 382 PHOSPHORYLATION (BY SIMILARITY).

FT BINDING 81 81 ASSOCIATED WITH AMINOACYL-TRNA BINDING  
 (BY SIMILARITY).  
 FT BINDING 137 137 ASSOCIATED WITH GUANOSINE NUCLEOTIDE  
 BINDING ACTIVITIES (BY SIMILARITY).  
 SQ SEQUENCE 393 AA; 43152 MW; 2833061693363BC4 CRC64;  
 Query Match 90.2%; Score 111; DB 1; Length 393;  
 Best Local Similarity 90.5%; Pred. No. 8.6e-11;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 KEKFNRTKPHVNIQIGHVDH 22  
 ||||| ||||| ||||| |||||  
 Db 2 KEKERTKPHVNVGTIGHVDH 22  
 RESULT 6  
 EFTU\_PASMU  
 ID EFTU\_PASMU STANDARD; PRT; 394 AA.  
 AC P57939;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Elongation factor Tu-A (EF-Tu-A).  
 GN TUPA OR PM1357.  
 OS Pasteurella multocida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Pasteurella.  
 ON NCBI\_TaxID=747;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM70;  
 RX MEDLINE=21145866; PubMed=11248100;  
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
 RT "Complete genomic sequence of *Pasteurella multocida* PM70.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 BIOSYNTHESIS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE006173; AAK03441.1; -;  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFACFOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding;  
 KW Complete proteome.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 81 85 GTP (BY SIMILARITY).  
 FT NP\_BIND 136 139 GTP (BY SIMILARITY).  
 SQ SEQUENCE 394 AA; 43353 MW; 0C704405CF5E9A4C CRC64;  
 Query Match 90.2%; Score 111; DB 1; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 8.6e-11;  
 Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;





```

RC STRAIN=NIES-484;
RX MEDLINE=99398694; PubMed=10468594;
RA Turmel M., Otis C., Lemieux C.;
RT "The complete chloroplast DNA sequence of the green alga Nephroselmis
RT olivacea: Insights into the architecture of ancestral chloroplast
RT genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:10248-10253(1999).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF137379; AAD54821.1; -.
CC HSP; P02990; 1ETU.
CC InterPro: IPR000795; GTP_EFTU.
CC Pfam; PF00009; GTP_EFTU; 1.
CC PRINTS; PR00315; ELONGATNFCT.
CC PROSITE; PS00301; EFACOR.GTP; 1.
CC Elongation factor; Protein biosynthesis; Chloroplast;
CC GTP-binding.
CC NP_BIND 19 26 GTP (BY SIMILARITY).
CC NP_BIND 81 85 GTP (BY SIMILARITY).
CC NP_BIND 136 139 GTP (BY SIMILARITY).
CC SEQUENCE 410 AA; 44810 MW; A7A9B0A6E99C1B23 CRC64;

Query Match 88.6%; Score 109; DB 1; Length 410;
Best Local Similarity 90.5%; Pred. No. 1.9e-10;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTGTTIGHVDH 22
Db :||| ||||| ||||| ||||| |||||
3 KEFERTKPHVNTGTTIGHVDH 23

RESULT 10
EFTU_CAMJE STANDARD; PRT; 399 AA.
AC O69303;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR TUFB OR CJ0470.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RA Thies F., Giegerich G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Whitt B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajadaram M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
```

```

RL Nature 403:665-668(2000).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y17167; CAAY66676.1; -.
CC HSP; ALI39075; CAB75108.1; -.
CC HSP; P02990; 1EFU.
CC InterPro: IPR000795; GTP_EFTU.
CC InterPro: IPR004161; GTP_EFTU_D2.
CC InterPro: IPR004160; GTP_EFTU_D3.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC Pfam; PF03143; GTP_EFTU_D3; 1.
CC PRINTS; PR00315; ELONGATNFCT.
CC PROSITE; PS00301; EFACOR.GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
CC Complete proteome.
CC NP_BIND 19 26 GTP (BY SIMILARITY).
CC NP_BIND 81 85 GTP (BY SIMILARITY).
CC NP_BIND 136 139 GTP (BY SIMILARITY).
CC SEQUENCE 399 AA; 43593 MW; OFAF86F46938AAA9 CRC64;

Query Match 87.8%; Score 108; DB 1; Length 399;
Best Local Similarity 90.5%; Pred. No. 2.7e-10;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNTGTTIGHVDH 22
Db :||| ||||| ||||| ||||| |||||
3 KEFERNKPHVNTGTTIGHVDH 23

RESULT 11
EFTU_NEIGO STANDARD; PRT; 394 AA.
AC P48864;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR TUFA.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX MEDLINE=96425873; PubMed=8828215;
RA Porcella S.F., Belland R.J., Judd R.C.;
RT "Identification of an EF-Tu protein that is periplasm-associated and
RT processed in Neisseria gonorrhoeae.";
RL Microbiology 142:2481-2489(1996).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; L36380; AAB41517.2; -.  
 DR HSSP; P02990; LETU.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFACOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 81 85 GTP (BY SIMILARITY).  
 FT NP\_BIND 136 139 GTP (BY SIMILARITY).  
 SQ SEQUENCE 394 AA; 43084 MW; F96799BE63520444 CRC64;

Query Match 87.0%; Score 107; DB 1; Length 394;  
 Best Local Similarity 85.7%; Pred. No. 3.9e-10;  
 Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNTRKPHVNIQTIGHVDH 22  
 |||||:|||||:|||||  
 DB 3 KEKFNTRKPHVNIQTIGHVDH 23

## RESULT 12

EFTU\_RECAM STANDARD; PRT; 394 AA.  
 AC Q21245;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Elongation factor Tu, mitochondrial.  
 GN TUFA.  
 OS Reclinomonas americana.  
 OG Mitochondrion.  
 OC Eukaryota; core jakobids; Reclinomonas.  
 OX NCBI\_TaxID=48483;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 50394;  
 RX MEDLINE=97311393; PubMed=9168110;  
 RA Lang B.F., Burger G., O'Kelly C.J., Cedergren R., Golding G.B.,  
 RA Lemieux C., Sankoff D., Turmel M., Gray M.W.;  
 RT "An ancestral mitochondrial DNA resembling a eubacterial genome in  
 RT miniature".  
 RL Nature 387:493-497(1997).  
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 CC BIOSYNTHESIS.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF007261; AAD11872.1; -.  
 DR HSSP; P02990; LETU.

DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFCT.  
 DR PROSITE; PS00301; EFACOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; Mitochondrion; GTP-binding.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 81 85 GTP (BY SIMILARITY).  
 FT NP\_BIND 136 139 GTP (BY SIMILARITY).  
 SQ SEQUENCE 394 AA; 43627 MW; 428202ADF3DBC4EF CRC64;

Query Match 87.0%; Score 107; DB 1; Length 394;  
 Best Local Similarity 90.5%; Pred. No. 3.9e-10;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNTRKPHVNIQTIGHVDH 22  
 |||||:|||||:|||||  
 DB 3 KEKFNTRKPHVNIQTIGHVDH 23

## RESULT 13

EFTU\_STIAU STANDARD; PRT; 396 AA.  
 AC P42479; Q53775;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Elongation factor Tu (EF-Tu).  
 GN TUF OR TUFB.  
 OS Stigmatella aurantiaca.  
 OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
 OC Myxococcales; Cystobacterineae; Cystobacteraceae; Stigmatella.  
 OX NCBI\_TaxID=41;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DW4;  
 RX MEDLINE=95303606; PubMed=7784178;  
 RA Bremaud L., Fremaux C., Laalami S., Genatiempo Y.;  
 RT "Genetic and molecular analysis of the trna-tufb operon of the  
 RT myxobacterium Stigmatella aurantiaca".  
 RL Nucleic Acids Res. 23:1737-1743(1995).  
 RN [2]  
 RP SEQUENCE OF 18-396 FROM N.A.  
 RC STRAIN=SGA1;  
 RX MEDLINE=94368062; PubMed=8085791;  
 RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Roller C.,  
 RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,  
 RA Bachleitner M., Fischer U., Schleifer K.H.;  
 RT "Phylogenetic relationships of Bacteria based on comparative sequence  
 RT analysis of elongation factor Tu and ATP-synthase beta-subunit  
 RT genes".  
 RL Antonie Van Leeuwenhoek 64:285-305(1993).  
 CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 CC BIOSYNTHESIS.  
 CC -!- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X82820; CAA58029.1; -.  
 DR HSSP; P02990; LETU.





```

CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL: AJ000260; CAA03976.1; -.
CC DR HSSP: P02990; IFTU.
CC DR InterPro: IPR000795; GTP_EFTU.
CC DR InterPro: IPR004161; GTP_EFTU_D2.
CC DR InterPro: IPR004160; GTP_EFTU_D3.
CC DR Pfam: PF00009; GTP_EFTU; 1.
CC DR Pfam: PF03144; GTP_EFTU_D2; 1.
CC DR Pfam: PF03143; GTP_EFTU_D3; 1.
CC DR PRINTS: PR00315; ELONGATNFCT.
CC DR PROSITE: PS00301; EFATOR_GTP; 1.
CC KW Elongation factor; Protein biosynthesis; GTP-binding.
CC FT NP_BIND 19 26 GTP (BY SIMILARITY).
CC FT NP_BIND 81 85 GTP (BY SIMILARITY).
CC FT NP_BIND 136 139 GTP (BY SIMILARITY).
CC SQ SEQUENCE 395 AA; 43290 MW; 71B08165E7FB42C5 CRC64;
CC -----
Query Match 86.2%; Score 106; DB 1; Length 395;
Best Local Similarity 90.5%; Pred. No. 5,8e-10;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNICTIGHVDH 22
   I I I I I I I I I I I I I I I I
Db 3 KAKPERTKPHVNICTIGHVDH 23

RESULT 18
EFTU_CORGL ID EFTU_CORGL STANDARD; PRT; 396 AA.
AC P42439;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OC NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13059 / AS019;
RX MEDLINE=94368062; PubMed=8085791;
RA Ludwig W., Neumaier J., Klugbauer N., Brockmeyer E., Røller C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,
RA Bachtelner M., Fischer U., Schleifer K.H.;
RT Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and ATP-synthase beta-subunit
RT genes.
RL Antonie Van Leeuwenhoek 64:285-305(1993).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN-TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RL "Massive gene decay in the leprosy bacillus.";
RT Nature 409:1007-1011(2001).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L13276; AAA71969.1; -.
DR EMBL; Z14314; CAA78674.1; -.
DR EMBL; D13869; BAA02982.2; -.
DR EMBL; AL583923; CAC30831.1; -.
DR PIR; S34954; S34954.
DR HSP; P02990; 1ETU.
DR Leproma; ML1877; -.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
DR Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
FT FT CONFLICT 129 152 GVPYLVALKNSDAVDDEELV -> VVLTWSHLSPT
FT FT CONFLICT 281 281 L -> V (IN REF. 2).
FT FT CONFLICT 349 349 G -> A (IN REF. 2).
FT FT CONFLICT 384 384 R -> P (IN REF. 3).
SQ SEQUENCE 396 AA; 43667 MW; D9CB88343C642778 CRC64;

Query Match 86.28; Score 106; DB 1; Length 396;
Best Local Similarity 90.58; Pred. No. 5.8e-10;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNICTIGHVDH 22
| | | | | | | | | | | | | | | | | | | | |
Db 3 KAKFERTKPHVNICTIGHVDH 23

RESULT 20
EFTU_MYCTU STANDARD; PRT; 396 AA.
AC P31501; Q50823; P95031;
DT 01-JUL-1993 (Rel. 26, Created)
DT* 01-JUL-1993 (Rel. 26, Last sequence update)

```

```

DT DE Elongation factor Tu (EF-Tu).
GN TUF OR RV0685 OR MT0713 OR MTCY210.02.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
OX Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN=ERDMANN;
RC MEDLINE=92347983; PubMed=1639483;
RA Carlin N.I.A., Loeffel S., Magnusson M.;
RT "Monoclonal antibodies specific for elongation factor Tu and complete
RL nucleotide sequence of the tuf gene in Mycobacterium tuberculosis.";
RN Infect. Immun. 60:3136-3142(1992).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence.";
RN Nature 393:537-544(1998).
[3]
RP SEQUENCE FROM N.A.
RX STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains.";
RE Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63539; CAA45102.1; -.
DR EMBL; X63539; CAA45101.1; ALT_INIT.
DR EMBL; Z84395; CAB06471.1; -.
DR EMBL; AE006965; AAK44939.1; -.
DR PIR; A44795; A44795.
DR HSP; P02990; 1ETU.
DR TIGR; MT0713; -.
DR TubercuList; Rv0685; -.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.

```

```

KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 396 AA; 43593 MW; 85F9012F364692FF CRC64;

Query Match 86.2%; Score 106; DB 1; Length 396;
Best Local Similarity 90.5%; Pred. No. 5.8e-10;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKENRTKPHVNICTIGHVDH 22
   | | | | | | | | | | | | | |
DB 3 KAKFORTKPHVNICTIGHVDH 23

RESULT 21
EFTU_THICU
ID EFTU_THICU STANDARD; PRT; 396 AA.
AC P42481; Q05066;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Thiobacillus cuprinus.
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Thiomonas.
OX NCBI_TaxID=36860;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HOE5;
RX MEDLINE=94368062; PubMed=8085791;
RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Rollier C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,
RA Bachleitner M., Fischer U., Schleifer K.H.;
RT "Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and ATP-synthase beta-subunit
RT genes.";
RL Antonie Van Leeuwenhoek 64:285-305(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 5495;
RA Moreira D., Amils R.;
RT cuprinus.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X76871; CAA54198.1; -;
CC EMBL; U78300; AAB87734.1; -;
CC HSSP; P02990; 1ETU.
CC InterPro; IPR000795; GTP_EFTU.
CC InterPro; IPR004161; GTP_EFTU_D2.
CC InterPro; IPR004160; GTP_EFTU_D3.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC Pfam; PF03143; GTP_EFTU_D3; 1.
CC PRINTS; PR00315; ELONGATNFCT.

```

DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding;  
 Complete proteome.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 81 85 GTP (BY SIMILARITY).  
 FT NP\_BIND 136 139 GTP (BY SIMILARITY).  
 SQ SEQUENCE 397 AA; 43369 MW; A019D5BF8EBAB942 CRC64;

Query Match 86.2%; Score 106; DB 1; Length 397;  
 Best Local Similarity 85.7%; Pred. No. 5.8e-10;  
 Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
 ||||| |||||:|||||  
 DB 3 KEKFNRTKPHVNIPTIGHVDH 23

## RESULT 23

EFTU\_STRCJ  
 ID EFTU\_STRCJ STANDARD; PRT; 397 AA.  
 AC P95724;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Elongation factor Tu (EF-Tu).  
 GN TUP.  
 OS Streptomyces cinnamonensis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxID=53446;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TUE89;  
 RA Cappellano C., Monti F., Sosio M., Donadio S., Sarubbi E.;  
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 CC BIOSYNTHESIS.  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----

DR EMBL; X98831; CAA67349.1; -  
 DR HSP; P02990; 1FTU  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 83 87 GTP (BY SIMILARITY).  
 FT NP\_BIND 138 141 GTP (BY SIMILARITY).  
 SQ SEQUENCE 397 AA; 43891 MW; EEB8235730D8BBB7 CRC64;

Query Match 86.2%; Score 106; DB 1; Length 397;  
 Best Local Similarity 90.5%; Pred. No. 5.8e-10;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
 ||||| |||||:|||||  
 DB 3 KAKFERTKPHVNIPTIGHVDH 23

## RESULT 24

EFTU\_STROR  
 ID EFTU\_STROR STANDARD; PRT; 398 AA.  
 AC P33170;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Elongation factor Tu (EF-Tu).  
 GN TUP.  
 OS Streptococcus oralis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90240875; PubMed-2110445;  
 RA Ludwig W., Weizenegger M., Betzl D., Leidel E., Lenz T., Ludvigsen A.,  
 RA Moellenhoff D., Wenzig P., Schleifer K.H.;  
 RT \*Complete nucleotide sequences of seven eubacterial genes coding for  
 RT the elongation factor Tu: functional, structural and phylogenetic  
 RT evaluations.\*;  
 RL Arch. Microbiol. 153:241-247(1990).  
 CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF  
 CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN  
 CC BIOSYNTHESIS.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC EF-TU/EF-1A SUBFAMILY.  
 DR PIR; F60663; F60663.  
 DR HSP; P02990; 1FTU.  
 DR InterPro; IPR000795; GTP\_EFTU.  
 DR InterPro; IPR004161; GTP\_EFTU\_D2.  
 DR InterPro; IPR004160; GTP\_EFTU\_D3.  
 DR Pfam; PF00009; GTP\_EFTU; 1.  
 DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
 DR Pfam; PF03143; GTP\_EFTU\_D3; 1.  
 DR PRINTS; PR00315; ELONGATNFT.  
 DR PROSITE; PS00301; EFATOR\_GTP; 1.  
 KW Elongation factor; Protein biosynthesis; GTP-binding.  
 FT NP\_BIND 19 26 GTP (BY SIMILARITY).  
 FT NP\_BIND 84 88 GTP (BY SIMILARITY).  
 FT NP\_BIND 139 142 GTP (BY SIMILARITY).  
 SQ SEQUENCE 398 AA; 44013 MW; 82399826C6C08E90 CRC64;

Query Match 86.2%; Score 106; DB 1; Length 398;  
 Best Local Similarity 85.7%; Pred. No. 5.8e-10;  
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIPTIGHVDH 22  
 ||||| |||||:|||||  
 DB 3 KEKYDRSKPHVNIPTIGHVDH 23

## RESULT 25

EFTU\_STRPY  
 ID EFTU\_STRPY STANDARD; PRT; 398 AA.  
 AC P82559;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)



```

DE Elongation factor Tu (EF-Tu).
GN TUF OR SPY0611.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [2]
RP PARTIAL SEQUENCE, AND MASS SPECTROMETRY.
RC STRAIN-JRS4 / Serotype M6;
RX Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes
RT proteins.";
RL Submitted (MAY-2000) to the SWISS-PROT data bank.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AE006516; AAK33586.1; -
CC InterPro: IPR000795; GTP_EFTU.
CC InterPro: IPR004161; GTP_EFTU_D2.
CC InterPro: IPR004160; GTP_EFTU_D3.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC Pfam: PF03143; GTP_EFTU_D3; 1.
CC PRINTS: PR00315; ELONGATNFCT.
CC PROSITE: PS00301; EFACOR_GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
CC Complete proteome.
CC NP_BIND 19 26 GTP (BY SIMILARITY).
CC NP_BIND 84 88 GTP (BY SIMILARITY).
CC NP_BIND 139 142 GTP (BY SIMILARITY).
CC SEQUENCE 398 AA; 43855 MW; 21E0B20A8DD21AFB CRC64;

Query Match 86.2%; Score 106; DB 1; Length 398;
Best Local Similarity 85.7%; Pred. No. 5.8e-10;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Oy 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KEKYDRSKPHVNIQIGHVDH 23
||||:|||||

RESULT 26
EFTU_RHILO
ID EFTU_RHILO STANDARD; PRT; 391 AA.
AC Q981F7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Elongation factor Tu (EF-Tu).
GN (TUF OR MLR0263) AND (TUFB OR MLR0288).
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: AP002994; BAB47886.1; -
CC EMBL: AP002994; BAB47904.1; -
CC InterPro: IPR000795; GTP_EFTU.
CC InterPro: IPR004161; GTP_EFTU_D2.
CC InterPro: IPR004160; GTP_EFTU_D3.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC Pfam: PF03143; GTP_EFTU_D3; 1.
CC PRINTS: PR00315; ELONGATNFCT.
CC PROSITE: PS00301; EFACOR_GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
CC Complete proteome.
CC NP_BIND 19 26 GTP (BY SIMILARITY).
CC NP_BIND 76 80 GTP (BY SIMILARITY).
CC NP_BIND 131 134 GTP (BY SIMILARITY).
CC SEQUENCE 391 AA; 42713 MW; 036718E6A48B48A9 CRC64;

Query Match 85.4%; Score 105; DB 1; Length 391;
Best Local Similarity 90.5%; Pred. No. 8.4e-10;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 KEKFNRTKPHVNIQIGHVDH 22
Db 3 KGKFNRTKPHVNIQIGHVDH 23
||||:|||||

RESULT 27
EFTU_MYCGA
ID EFTU_MYCGA STANDARD; PRT; 394 AA.
AC P18906;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Mycoplasma gallisepticum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.

```

```
OX NCBI_TaxID=2096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098800; PubMed=2602129;
RA Inamine J.M., Loechel S., Hu P.C.;
RT "Nucleotide sequence of the tuf gene from Mycoplasma gallisepticum.";
RL Nucleic Acids Res. 17:10126-10126(1989).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X16462; CAA34482.1; -.
CC PIR; S14910; EFMYS.
CC HSP; P02990; 1EFU.
CC InterPro; IPR000795; GTP_EFTU.
CC InterPro; IPR004161; GTP_EFTU_D2.
CC InterPro; IPR004160; GTP_EFTU_D3.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC Pfam; PF03143; GTP_EFTU_D3; 1.
CC PRINTS; PR00315; ELONGATNFCT.
CC PROSITE; PS00301; EFATOR.GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding.
KW NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 394 AA; 43099 MW; 963E8B93216279F1 CRC64;

Query Match 85.4%; Score 105; DB 1; Length 394;
Best Local Similarity 81.0%; Pred. No. 8.4e-10;
Matches 17; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNICTIGHVDH 22
   ||:|||||:|||||:|
Db 3 KERFDRSKPHVNICTIGHIDH 23

RESULT 28
EFTL_STRCO STANDARD; PRT; 397 AA.
AC P40174;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tuf-1 (EF-Tu-1).
GN TUF1 OR SCD40A.08.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RX MEDLINE=95002174; PubMed=7918656;
RA van Wezel G.P., Woudt L.P., Vervenne R., Verdurmen M.L.,
RA Vrijboom E., Bosch L.;
RT "Cloning and sequencing of the tuf genes of Streptomyces coelicolor
RT A3(2).";
RL Biochim. Biophys. Acta 1219:543-547(1994).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,
RA Rajandream M.A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X77039; CAA54329.1; -.
CC PIR; AL161691; CAB81853.1; -.
CC HSP; P02990; 1EFU.
CC InterPro; IPR000795; GTP_EFTU.
CC InterPro; IPR004161; GTP_EFTU_D2.
CC InterPro; IPR004160; GTP_EFTU_D3.
CC Pfam; PF00009; GTP_EFTU; 1.
CC Pfam; PF03144; GTP_EFTU_D2; 1.
CC Pfam; PF03143; GTP_EFTU_D3; 1.
CC PRINTS; PR00315; ELONGATNFCT.
CC PROSITE; PS00301; EFATOR.GTP; 1.
CC Elongation factor; Protein biosynthesis; GTP-binding;
KW Multigene family.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 83 87 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 397 AA; 43781 MW; D27DF957DF1DDF75 CRC64;

Query Match 85.4%; Score 105; DB 1; Length 397;
Best Local Similarity 85.7%; Pred. No. 8.5e-10;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNICTIGHVDH 22
   ||:|||||:|||||:|
Db 3 KAKFERTKPHVNICTIGHIDH 23

RESULT 29
EFTL_STRCU STANDARD; PRT; 397 AA.
AC Q53871;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tuf-1 (EF-Tu-1).
GN TUF1
OS Streptomyces collinus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=42684;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BSM 40733;
RX MEDLINE=95374503; PubMed=7646499;
RA Mikulik K., Zhulanova E.;
RT "Sequencing of the tuf1 gene and the phosphorylation pattern of
RT EF-Tu1 during development and differentiation in Streptomyces
RT collinus producing kirromycin.";
RL Biochem. Biophys. Res. Commun. 213:454-461(1995).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
```



```

DR DR EMBL1: AF007125; NAB62702.1; -.
DR DR HSSP: P02990; 1FTU.
DR DR InterPro: IPR000795; GTP_EFTU.
DR DR InterPro: IPR004161; GTP_EFTU_D2.
DR DR InterPro: IPR004160; GTP_EFTU_D3.
DR DR pfam: PF00009; GTP_EFTU; 1.
DR DR pfam: PF03144; GTP_EFTU_D2; 1.
DR DR pfam: PF03143; GTP_EFTU_D3; 1.
DR DR PRINTS: PR00315; ELONGINFCT.
DR DR PROSITE: PS00301; EFAPTOR_GTP; 1.

```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-si.ch/announce/>)

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z11874; CAA77904.1; -
DR EMBL; X0044; CAA24925.1; -
DR EMBL; X06254; CAA29599.1; -
DR EMBL; X70810; CAA50087.1; -
DR EMBL; X00480; CAA25159.1; -
DR PIR; A03519; EFEGT.
DR PIR; S02254; S02254.
DR PIR; S34508; S34508.
DR HSSP; P02990; LETU.
DR Mendel; 4838; EUGgr; Tufa.1.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU.1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; Chloroplast;
KW GTP-binding; Methylation.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
FT MOD_RES 57 57 METHYLATION (MONO-) (PROBABLE).
SQ SEQUENCE 409 AA; 45062 MW; C803740422FFEB84 CRC64;

Query Match 85.4%; Score 105; DB 1; Length 409;
Best Local Similarity 81.0%; Pred. No. 8.8e-10;
Matches 17; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNTPKPHVNIQTIGHVDH 22
DB 3 ROKFERTKPHINVTIGHVDH 23

RESULT 34
EFTU_BUCAI
ID EFTU_BUCAI STANDARD; PRT; 394 AA.
AC O31297;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR BU526
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
symbolic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TOKYO 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RY Brynne E.U., Kurland C.G., Moran N.A., Andersson S.G.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS."
RL Nature 407:81-86(2000).
RN [2]
RP SEQUENCE OF 20-384 FROM N.A.
RX MEDLINE=98242088; PubMed=9580987;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RY Brynne E.U., Kurland C.G., Moran N.A., Andersson S.G.;
RT "Evolutionary rates for tuf genes in endosymbionts of aphids."
RL Mol. Biol. Evol. 15:574-582(1998).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001119; BAB13219.1; ALT_INIT.
DR EMBL; Y12307; CAA72974.1; -
DR HSSP; P02990; LETU.
DR InterPro; IPR000795; GTP_EFTU.
DR Pfam; PF00009; GTP_EFTU.1.
DR PRINTS; PR00315; ELONGATNFCT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
FT CONFLICT 89 89 I -> M (IN REF. 2).
SQ SEQUENCE 394 AA; 43465 MW; 09B73EADCA0DF5F6 CRC64;

Query Match 84.6%; Score 104; DB 1; Length 394;
Best Local Similarity 81.0%; Pred. No. 1.2e-09;
Matches 17; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KEKFNTPKPHVNIQTIGHVDH 22
DB 3 KEKFNTPKPHVNIQTIGHVDH 23

RESULT 35
EFTU_MYCGE
ID EFTU_MYCGE STANDARD; PRT; 394 AA.
AC P13927; Q49360;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR MG451.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=90098801; PubMed=26021130;
RA Inamine J.M., Loechel S., Hu P.-C.;
RT "Nucleotide sequence of the tuf gene from Mycoplasma genitalium."
RL Nucleic Acids Res. 17:10127-10127(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium."
RL Science 270:337-403(1995).
RN [3]
RP SEQUENCE OF 196-309 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
sequencing."
RL J. Bacteriol. 175:7918-7930(1993).
```



```

RX MEDLINE=90240875; PubMed=2110445;
RA Ludwig W., Weizenegger M., Betzl D., Leidel E., Lenz T., Ludvigsen A.,
RA Moellenhoff D., Wenzig P., Schleifer K.H.;
RT *Complete nucleotide sequences of seven eubacterial genes coding for
RT the elongation factor Tu: functional, structural and phylogenetic
RL evaluations.*;
RL Arch. Microbiol. 153:241-247(1990).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC PTR: D60663; D60663.
CC DR HSSP; P02990; 1FTU.
CC DR InterPro; IPR000795; GTP_EFTU.
CC DR InterPro; IPR004161; GTP_EFTU_D2.
CC DR InterPro; IPR004160; GTP_EFTU_D3.
CC DR Pfam; PF00009; GTP_EFTU; 1.
CC DR Pfam; PF03144; GTP_EFTU_D2; 1.
CC DR Pfam; PF03143; GTP_EFTU_D3; 1.
CC DR PRINTS; PR00315; ELONGATNFCT.
CC DR PROSITE; PS00301; EFATOR_GTP; 1.
CC KW Elongation factor; Protein biosynthesis; GTP-binding.
CC FT NP_BIND 19 26 GTP (BY SIMILARITY).
CC FT NP_BIND 81 85 GTP (BY SIMILARITY).
CC FT NP_BIND 136 139 GTP (BY SIMILARITY).
CC SQ SEQUENCE 396 AA; 42876 MW; EE21378647AFD644 CRC64;

Query Match 84.6%; Score 104; DB 1; Length 396;
Best Local Similarity 85.7%; Pred. No. 1.2e-09;
Matches 18; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 KGKFKRTKPHVNVGTIGHVDH 23

RESULT 38
EFTU_HELPJ
ID EFTU_HELPJ STANDARD; PRT; 399 AA.
AC Q9ZK19;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF OR TUFU OR JHP1128.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT *Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.*;
RL Nature 397:176-180(1999).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001541; AAD06711.1; -.
CC HSSP; P02990; 1FTU.
CC DR InterPro; IPR000795; GTP_EFTU.
CC DR InterPro; IPR004161; GTP_EFTU_D2.
CC DR InterPro; IPR004160; GTP_EFTU_D3.
CC DR Pfam; PF00009; GTP_EFTU; 1.
CC DR Pfam; PF03144; GTP_EFTU_D2; 1.
CC DR Pfam; PF03143; GTP_EFTU_D3; 1.
CC DR PRINTS; PR00315; ELONGATNFCT.
CC DR PROSITE; PS00301; EFATOR_GTP; 1.
CC KW Elongation factor; Protein biosynthesis; GTP-binding;
CC KW Complete proteome.
CC FT NP_BIND 19 26 GTP (BY SIMILARITY).
CC FT NP_BIND 81 85 GTP (BY SIMILARITY).
CC FT NP_BIND 136 139 GTP (BY SIMILARITY).
CC SQ SEQUENCE 399 AA; 43730 MW; 4E72A877BFCD104B CRC64;

Query Match 84.6%; Score 104; DB 1; Length 399;
Best Local Similarity 90.5%; Pred. No. 1.3e-09;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KEKFNRTKPHVNIQTIGHVDH 22
Db 3 KEKFNRTNPHVNIQTIGHVYH 23

RESULT 39
EFTU_TAXOC
ID EFTU_TAXOC STANDARD; PRT; 395 AA.
AC P42480;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Taxeobacter ocellatus.
OC Bacteria; CFB group; Taxeobacter.
OX NCBI_TaxID=36878;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MYX 2105;
RX MEDLINE=94368062; PubMed=8085791;
RA Ludwig W., Neumaier J., Klugbauer N., Brockmann E., Roller C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,
RA Bachleitner M., Fischer U., Schleifer K.H.;
RT *Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and Arp-synthase beta-subunit
RT genes.*;
RL Antonie Van Leeuwenhoek 64:285-305(1993).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; X77036; CAA54325.1; -.
DR HSSP; P02990; IEFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR Pfam; PF03143; GTP_EFTU_D3; 1.
DR PRINTS; PS00315; ELONGATNFACT.
DR PROSITE; PS00301; EFATOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 81 85 GTP (BY SIMILARITY).
FT NP_BIND 136 139 GTP (BY SIMILARITY).
SQ SEQUENCE 395 AA; 43038 MW; 168222411386D156 CRC64;

Query Match 83.7%; Score 103; DB 1; Length 395;
Best Local Similarity 85.7%; Pred. No. 1.8e-09;
Matches 18; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KEKFNRTKPHVNIQTIGHVDH 22
   ||| |.|||||
DB 3 KEFDKSKPHVNIQTIGHVDH 23

RESULT 40
EFTU_HERAU
ID EFTU_HERAU STANDARD; PRT; 400 AA.
AC P42477;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Elongation factor Tu (EF-Tu).
GN TUF.
OS Herpetosiphon aurantiacus (Herpetosiphon giganteus).
OC Bacteria; Green non-sulfur bacteria; Chloroflexaceae group;
OC Herpetosiphon.
OX NCBI_TaxID=65;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HPGAL;
RX MEDLINE=94388062; PubMed=8085791;
RA Ludwig W., Neumaier J., Kludbauer N., Brockmann E., Roller C.,
RA Klugbauer S., Reetz K., Schachtner I., Ludvigsen A.,
RA Bachleitner M., Fischer U., Schleifer K.H.;
RT "Phylogenetic relationships of Bacteria based on comparative sequence
RT analysis of elongation factor Tu and ATP-synthase beta-subunit
RT genes.";
RL Antonie van Leeuwenhoek 64:285-305(1993).
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC BIOSYNTHESIS.
CC -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-TU/EF-1A SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X76868; CAA54196.1; -.
DR HSSP; P02990; IEFU.
DR InterPro; IPR000795; GTP_EFTU.
DR InterPro; IPR004161; GTP_EFTU_D2.
DR InterPro; IPR004160; GTP_EFTU_D3.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:45:38 ; Search time 13.53 Seconds

(without alignments)  
82.991 Million cell updates/sec

Title: US-09-488-737-2

Perfect score: 157

Sequence: 1 MNKVDVKQTAFGAPWDDNNVITAGPRG 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match Length	DB ID	Description
1	157	100.0	505	1	CATA_HELPJ
2	157	100.0	505	1	CATA_HELPJ
3	93	59.2	484	1	CATA_PROMI
4	88	56.1	482	1	CATA_BORPE
5	83.5	53.2	491	1	CATA_ORYSA
6	82	52.2	496	1	CAT3_MAIZE
7	78.5	50.0	494	1	CAT2_HORVU
8	78	49.7	482	1	CATA_PSEAE
9	77.5	49.4	492	1	CATA_SECE
10	76	48.4	492	1	CAT1_GOSHI
11	75	47.8	492	1	CAT1_CUCPE
12	75	47.8	492	1	CAT2_ARATH
13	74	47.1	492	1	CAT1_HORVU
14	74	47.1	492	1	CAT2_WHEAT
15	73	46.5	479	1	CATA_DESVM
16	73	46.5	492	1	CATA_PHAU
17	72	45.9	505	1	CATA_METBA
18	71	45.2	482	1	CATA_BACSU
19	71	45.2	494	1	CATA_RHIME
20	70	44.6	492	1	CAT1_MAIZE
21	70	44.6	492	1	CAT1_ORYSA
22	68	43.3	492	1	CAT1_ARATH
23	68	43.3	492	1	CAT1_RICCO
24	68	43.3	492	1	CAT2_GOSHI
25	68	43.3	492	1	CATA_SOLAP
26	68	43.3	494	1	CATA_PEA
27	67	42.7	479	1	CATA_PSEPU
28	67	42.7	492	1	CAT1_SOYBN
29	67	42.7	492	1	CAT3_NICPL
30	67	42.7	492	1	CAT3_SOYBN
31	67	42.7	492	1	CAT4_SOYBN
32	66	42.0	502	1	CATA_PIG
33	66	42.0	526	1	CATA_HUMAN

34	65	41.4	486	1	CATA_BACFR
35	65	41.4	508	1	CATA_HAIN
36	65	41.4	526	1	CATA_CANFA
37	65	41.4	526	1	CATA_CANFA
38	64	40.8	492	1	CAT1_WHEAT
39	64	40.8	492	1	CATA_HELAN
40	64	40.8	526	1	CATA_BRARE
41	64	40.8	526	1	CATA_RAT
42	63	40.1	491	1	CAT2_MAIZE
43	63	40.1	498	1	CATA_BRUME
44	62	39.5	482	1	CATA_ONCVE
45	62	39.5	705	1	CATE_RHIME
46	61	38.9	482	1	CATA_VIBFI
47	61	38.9	563	1	NOX1_RAT
48	60	38.2	492	1	CATA_IPOBA
49	60	38.2	502	1	CATA_TOXGO
50	59	37.6	492	1	CAT2_CUCPE
51	59	37.6	492	1	CAT3_CUCPE
52	59	37.6	527	1	CATA_RANRU
53	58	36.9	485	1	CAT1_NICPL
54	58	36.9	492	1	CAT2_LYCES
55	58	36.9	564	1	NOX1_HUMAN
56	57	36.3	339	1	REC7_SCHPO
57	57	36.3	492	1	CATA_AVIMR
58	57	36.3	496	1	CAT1_CABEL
59	57	36.3	506	1	CATA_BOVIN
60	56	35.7	500	1	CATA_NEIGO
61	56	35.7	526	1	CATA_MOUSE
62	55	35.7	541	1	CATA_ASCSU
63	55	35.0	492	1	CAT2_RICCO
64	55	35.0	513	1	CATB_PSEAE
65	54	34.4	492	1	CAT1_LYCES
66	54	34.4	492	1	CAT1_SOLTU
67	54	34.4	492	1	CAT1_TOBAC
68	54	34.4	492	1	CAT2_NICPL
69	54	34.4	492	1	CAT2_SOLTU
70	54	34.4	492	1	CATA_CAPAN
71	54	34.4	492	1	CATA_SOLME
72	54	34.4	784	1	OSTA_ECOLI
73	53	33.8	169	1	BENB_ACICA
74	52.5	33.4	730	1	CATR_ASPNG
75	52	33.1	488	1	CATA_LISSE
76	51	32.5	488	1	CATA_LISSE
77	50	31.8	483	1	BCA_STRVL
78	49	31.2	456	1	YC88_MYCTU
79	49	31.2	556	1	ASNS_SCHPO
80	49	31.2	576	1	LC1B_HUMAN
81	48	30.6	354	1	WN8D_MOUSE
82	48	30.6	355	1	WN8D_MOUSE
83	48	30.6	536	1	CATA_DETRA
84	48	30.6	547	1	CATX_BACSU
85	48	30.6	763	1	GLH1_CABEL
86	48	30.6	972	1	ORC4_SCHPO
87	47.5	30.3	728	1	CATB_ASPFU
88	47	29.9	307	1	YIDL_ECOLI
89	47	29.9	487	1	CATA_STRCO
90	47	29.9	617	1	YACH_ECOLI
91	47	29.9	623	1	PBP2_SALTU
92	46.5	29.6	570	1	C24B_BOVIN
93	46.5	29.6	753	1	ORCE_ECOLI
94	46	29.3	433	1	ORC4_MOUSE
95	46	29.3	468	1	YQK1_BACSU
96	46	29.3	484	1	C24B_PIG
97	46	29.3	569	1	C24B_HUMAN
98	46	29.3	570	1	C24B_MOUSE
99	46	29.3	648	1	TALA_POVWK
100	45.5	29.0	455	1	UFOG_HORVU

ALIGNMENTS

RESULT 1

```

CATA_HELPJ
ID CATA_HELPJ STANDARD; PRT; 505 AA.
AC Q9ZKX5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KATA OR JHP0809.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001510; AAD06391.1; -.
CC HSSP; P42321; 2CAE.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Complete proteome.
FT ACT_SITE 56 56 BY SIMILARITY.
FT ACT_SITE 129 129 BY SIMILARITY.
FT BINDING 339 339 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT BINDING 339 339 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 82 82 S -> F (IN REF. 1).
FT CONFLICT 234 234 V -> I (IN REF. 2).
FT CONFLICT 237 237 Y -> H (IN REF. 2).
FT CONFLICT 248 248 N -> D (IN REF. 2).
FT CONFLICT 255 255 F -> Y (IN REF. 2).
FT CONFLICT 286 286 L -> T (IN REF. 2).
FT CONFLICT 313 313 A -> V (IN REF. 1).

Query Match 100.0%; Score 157; DB 1; Length 505;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNKKVQTAFAGVWDDNNVITAGPRG 29
Db 1 MVNKKVQTAFAGVWDDNNVITAGPRG 29

RESULT 2
CATA_HELPJ
ID CATA_HELPJ STANDARD; PRT; 505 AA.
AC P77872; P94823;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KATA OR HP0875.
OS Helicobacter pylori (Campylobacter pylori).

```

```

OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98207782; PubMed=9546115;
RA Manos J., Kolesnikow T., Hazell S.L.;
RT "An investigation of the molecular basis of the spontaneous
RT occurrence of a catalase-negative phenotype in Helicobacter pylori.";
RL Helicobacter 3:28-38(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97113460; PubMed=8955320;
RA Odenbreit S., Wieland B., Haas R.;
RT "Cloning and genetic characterization of Helicobacter pylori catalase
RT and construction of a catalase-deficient mutant strain.";
RL J. Bacteriol. 178:6960-6967(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori.";
RL Nature 388:539-547(1997).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67458; AAC16068.1; -.
CC EMBL; Z70679; CAA94567.1; -.
CC EMBL; AF000597; AAD07923.1; -.
CC HSSP; P42321; 2CAE.
CC TIGR; HP0875; -.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Complete proteome.
FT ACT_SITE 56 56 BY SIMILARITY.
FT ACT_SITE 129 129 BY SIMILARITY.
FT BINDING 339 339 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT BINDING 339 339 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 82 82 S -> F (IN REF. 1).
FT CONFLICT 234 234 V -> I (IN REF. 2).
FT CONFLICT 237 237 Y -> H (IN REF. 2).
FT CONFLICT 248 248 N -> D (IN REF. 2).
FT CONFLICT 255 255 F -> Y (IN REF. 2).
FT CONFLICT 286 286 L -> T (IN REF. 2).
FT CONFLICT 313 313 A -> V (IN REF. 1).

```

FT	CONFLICT	316	316	S -> T (IN REF. 2).
SQ	SEQUENCE	505 AA;	58629 MW;	9F029B55B73C26EA CRC64;

SQ SEQUENCE 505 AA; 58629 MW; 9F029B55B73C26EA CRC64;

Query Match	Score 157:	DB 1:	Length 505:
100.08:			

Query Match 100.0%; Score 157; DB 1;

Best Local Similarity	100.0%;	Pred. No. 4.3e-16;	
Matches	29; Conservative	0; Mismatches	0; Indels
			0; Gaps

QY 1 MVNKDVKQTTAFGAPVWDDNNVITAGPRG 29

\_\_\_\_\_

### RESULT 3

CATA_PROMI	CATA_PROMI
ID	STANDARD;
	PRT; 484 AA.

AC P42321;

DT 01-NOV-2001

DT	01-NOV-1995	(Rel. 32,	Last seq
DT	01-NOV-1995	(Rel. 32,	Last seq

OS *Oryza sativa* (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoidae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A. TISSUE=Immature seed;  
 RC STRAIN=CV. INDICA; TISSUE=Leaf;  
 RX MEDLINE=92256818; PubMed=1581574;  
 RA Mori H., Higo K., Higo H., Minobe Y., Matsui H., Chiba S.;  
 RT "Nucleotide and derived amino acid sequence of a catalase cDNA  
 RT isolated from rice immature seeds.";  
 RL Plant Mol. Biol. 18:973-976(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. JAPONICA; TISSUE=Leaf;  
 RX MEDLINE=96189265; PubMed=8605302;  
 RA Higo K., Higo H.;  
 RT "Cloning and characterization of the rice Catalase gene, a  
 RT homologue of the maize Cat3 gene.";  
 RL Plant Mol. Biol. 30:505-521(1996).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE.  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -1- COFACTOR: HEME GROUP.  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X61626; CAA43814.1; -;  
 DR EMBL; D29966; BAA06232.1; -;  
 DR PIR; S20873; CSRZ.  
 DR HSSP; P00432; 4BLC.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Peroxisome; Glyoxysome; Multigene family.  
 FT ACT\_SITE 65 65 BY SIMILARITY.  
 FT ACT\_SITE 138 138 BY SIMILARITY.  
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 113 113 P -> O (IN REF. 2).  
 FT CONFLICT 448 449 AV -> PL (IN REF. 2).  
 SQ SEQUENCE 491 AA; 56606 MW; 5BA4CD73A9F4B748 CRC64;  
 Query Match 53.2%; Score 83.5; DB 1; Length 491;  
 Best Local Similarity 61.5%; Pred. No. 5.4e-05;  
 Matches 16; Conservative 2; Mismatches 7; Indels 1; Gaps 1;  
 QY 5 DVKQTTA-FGAPVWDDNNVTAGPRG 29  
 Db 13 DTKTTTNNAGAPVWDDNNVTAGPRG 38  
 RESULT 6  
 ID CAT3\_MAIZE STANDARD; PRT; 496 AA.  
 AC P18123;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 .DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Catalase isozyme 3 (EC 1.11.1.6).  
 GN CAT3.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoidae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. W64A; TISSUE=Seedling leaf;  
 RX MEDLINE=94003062; PubMed=8400123;  
 RA Adler M.L., Scandalios J.G.;  
 RT "Isolation and characterization of a genomic sequence encoding the  
 RT maize Cat3 catalase gene";  
 RL Plant Mol. Biol. 22:1031-1038(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. W64A; TISSUE=Epicotyl;  
 RX MEDLINE=89051000; PubMed=2461221;  
 RA Redinbaugh M.G., Wadsworth G.J., Scandalios J.G.;  
 RT "Characterization of catalase transcripts and their differential  
 RT expression in maize";  
 RL Biochim. Biophys. Acta 951:104-116(1988).  
 CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES  
 CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE. ITS  
 CC LEVELS ARE HIGHEST IN THE LIGHT PERIOD AND ARE LOWEST IN THE DARK  
 CC PERIOD, HENCE MAY BE IMPORTANT FOR SCAVENGING HYDROGEN PEROXIDE  
 CC AT NIGHT, THAN DURING THE DAY.  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -1- COFACTOR: HEME GROUP.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- TISSUE SPECIFICITY: LEAF MESOPHYLL CELLS, PERICARP, SEEDLING ROOTS  
 CC AND THE COLEOPTILE.  
 CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -1- CAUTION: THERE ARE TWO EMBL ENTRIES THAT CORRESPOND TO THE REF. 2  
 CC SEQUENCE AND THEY ENCODE SLIGHTLY DIFFERENT PROTEINS (SEE THE  
 CC FEATURE TABLE).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L05934; AAC37357.1; -;  
 DR EMBL; M33103; AAA33441.1; -;  
 DR EMBL; X12539; CAA31057.1; -;  
 DR PIR; JA0091; JA0091.  
 DR HSSP; P00432; 4BLC.  
 DR MaizeDB; 13855; -;  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Mitochondrion; Multigene family.  
 FT ACT\_SITE 67 67 BY SIMILARITY.  
 FT ACT\_SITE 140 140 BY SIMILARITY.  
 FT BINDING 351 351 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 57 57 A -> D (IN REF. 2).  
 FT CONFLICT 79 79 C -> S (IN REF. 2).  
 FT CONFLICT 107 107 T -> Q (IN REF. 2).  
 FT CONFLICT 110 110 H -> P (IN REF. 2).  
 FT CONFLICT 112 112 R -> P (IN REF. 2).  
 FT CONFLICT 115 116 PE -> GR (IN REF. 2).  
 FT CONFLICT 121 121 P -> A (IN REF. 2).  
 FT CONFLICT 194 194 D -> H (IN REF. 2; AAA33441).

FT CONFLICT 234 234 V -> E (IN REF. 2).

FT CONFLICT 236 236 C -> S (IN REF. 2; AAA33441).

FT CONFLICT 244 245 AL -> R (IN REF. 2).

FT CONFLICT 254 254 MISSING (IN REF. 2).

FT CONFLICT 264 264 A -> AE (IN REF. 2).

FT CONFLICT 280 281 DT -> AQ (IN REF. 2).

FT CONFLICT 282 282 E -> Q (IN REF. 2; AAA33441).

FT CONFLICT 319 319 F -> L (IN REF. 2).

FT CONFLICT 374 375 AH -> GT (IN REF. 2).

FT CONFLICT 386 386 F -> L (IN REF. 2).

FT CONFLICT 402 408 PLRQAP -> RRCRAA (IN REF. 2).

FT CONFLICT 452 465 REFADSLGHPKVSQ -> KAIRRLARTPORQP (IN REF. 2).

FT CONFLICT 478 478 C -> V (IN REF. 2).

FT CONFLICT 496 AA; 56796 MW; 566FFD05B3795B49 CRC64;

Query Match 52.2%; Score 82; DB 1; Length 496;

Best Local Similarity 66.7%; Pred. No. 9.2e-05;

Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29

DB 20 TTNAGAPVWDDNNVITAGPRG 40

II IIII:II:IIII

RESULT 7

CAT2\_HORVU STANDARD; PRT; 494 AA.

AC P55308;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Catalase isozyme 2 (EC 1.11.1.6).

GN CAT2.

OS Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;

OC Triticeae; Hordeum.

OX NCBI\_TaxID=4513;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=CV. MOREX;

RX MEDLINE=96145511; PubMed=8555444;

RA Skadsen R.W., Schulze-Lefert P., Herbst J.M.;

RT "Molecular cloning, characterization and expression analysis of two

RT catalase isozyme genes in barley.";

RL Plant Mol. Biol. 29:1005-1014(1995).

CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND

CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN

CC PEROXIDE.

CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.

CC -!- COFACTOR: HEME GROUP (BY SIMILARITY).

CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).

CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; U20778; AAA96948.1; -.

DR HSSP; P00432; 4BLC.

DR InterPro; IPR002226; Catalase.

DR Pfam; PF00199; catalase; 1.

DR PRINTS; PR00067; CATALASE.

DR ProDom; PD000510; Catalase; 1.

DR PROSITE; PS00437; CATALASE\_1; 1.

DR PROSITE; PS00438; CATALASE\_2; 1.

KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;

KW Peroxisome; Glyoxysome; Multigene family.

FT ACT\_SITE 65 65 BY SIMILARITY.

FT ACT\_SITE 138 138 BY SIMILARITY.

FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).

SQ SEQUENCE 494 AA; 56913 MW; 67E3CF2AD02E542E CRC64;

Query Match 50.0%; Score 78.5; DB 1; Length 494;

Best Local Similarity 57.7%; Pred. No. 0.00031;

Matches 15; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 5 DVKQTTA-FCGAPVWDDNNVITAGPRG 29

DB 13 DTKTTTNNAGQPWDDNNVITAGPRG 38

II IIII:II:IIII

RESULT 8

CATA\_PSEAE STANDARD; PRT; 482 AA.

AC O52762;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Catalase (EC 1.11.1.6).

GN KATA OR PA4236.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PRDJ.

RX MEDLINE=99296583; PubMed=10368148;

RA Ma J.-F., Ochsen U.A., Klotz M.G., Nanayakkara V.K., Howell M.L.,

RA Johnson Z., Posey J.E., Vasil M.L., Monaco J.J., Hassett D.J.;

RT "Bacterioferritin A modulates catalase A (Kata) activity and

RT resistance to hydrogen peroxide in Pseudomonas aeruginosa.";

RL J. Bacteriol. 181:3730-3742(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,

RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";

RL Nature 406:959-964(2000).

CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES

CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.

CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.

CC -!- COFACTOR: HEME GROUP.

CC -!- ENZYME REGULATION: BY PEROXIDE AND BFR-BOUND IRON.

CC -!- SUBUNIT: HETEROMULTIMER. POSSIBLY AN ALPHA(2)BETA-HETEROTRIMER

CC WHERE THE ALPHA SUBUNIT IS A 56 kDa PROTEIN AND THE BETA SUBUNIT A

CC 45 kDa PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AF047025; AAC03118.1; -.

DR EMBL; AF004841; AAG07624.1; -.

DR HSSP; P42321; 2CAE.

```

DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR ProDom: PD000510; Catalase; 1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Complete proteome.
FT ACT_SITE 55 55 BY SIMILARITY.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 338 338 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 482 AA; 55589 MW; 84E5ABA647CAB414 CRC64;

Query Match 49.7%; Score 78; DB 1; Length 482;
Best Local Similarity 57.1%; Pred. No. 0.00036;
Matches 16; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNKDVKQTAFGAPVWDDNNVITAGPRG 29
DB 1 MEEKTRLTAAAGAPVVDNQNVQTAGPRG 28
: : ||| ||| : || |||||

RESULT 9
CATA_SECE
ID CATA_SECE STANDARD; PRT; 492 AA.
AC P55310;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Catalase (EC 1.11.1.6).
OS Secale cereale (Rye).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Secale.
OX NCBI_TaxID=4550;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HALO; TISSUE=Leaf;
RA Schmidt M.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; Z541143; CAA90858.1; -.
DR HSP; P42321; 2CAE.
DR InterPro: IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome.
FT ACT_SITE 64 64 BY SIMILARITY.
FT ACT_SITE 137 137 BY SIMILARITY.
FT BINDING 347 347 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 492 AA; 56600 MW; D725C042259AE494 CRC64;

```

```

Query Match 49.4%; Score 77.5; DB 1; Length 492;
Best Local Similarity 57.7%; Pred. No. 0.00044;
Matches 15; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 5 DVKQTTAF-GAPVWDDNNVITAGPRG 29
DB 13 DTKTTTNGQPWWDNEALTVGPRG 38
: ||| ||| ||| ||| : |||||

RESULT 10
CAT1_GOSHI
ID CAT1_GOSHI STANDARD; PRT; 492 AA.
AC P17598;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE Catalase isozyme 1 (EC 1.11.1.6).
GN CAT1 OR SU1.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DELTAPINE 62; TISSUE=Cotyledon;
RX MEDLINE=90304227; PubMed=2364113;
RT Ni W., Turley R.B., Trelease R.N.;
RL "Characterization of a cDNA encoding cottonseed catalase.";
RL Biochim. Biophys. Acta 1049:219-222(1990).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: Peroxisomal.
CC -!- MISCELLANEOUS: THERE ARE AT LEAST FIVE ISOZYMES OF CATALASE IN
CC COTTON.
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X521135; CAA36380.1; -.
DR PIR; S10770; S10770.
DR PIR; S10395; S10395.
DR HSP; P00432; 4HLC.
DR InterPro: IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
DR Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 492 AA; 56855 MW; 6486966A095F261 CRC64;

Query Match 48.4%; Score 76; DB 1; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.00074;
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDDNNVITAGPRG 29

```

```

Db 18 TTNAGAPVNNSSSLTVGPRG 38
II IIIII::I::I::IIII
RESULT 11
CAT1_CUCPE STANDARD; PRT; 492 AA.
ID CAT1_CUCPE
AC P48350;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Catalase isozyme 1 (EC 1.11.1.6).
GN CAT1.
OS Cucurbita pepo (Vegetable marrow) (Summer squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3663;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cotyledon;
RX MEDLINE=97188581; PubMed=9037166;
RA Esaka M., Yamada N., Kitabayashi M., Setoguchi Y., Tsugeki R.,
RA Kondo M., Nishimura M.;
RT *cDNA cloning and differential gene expression of three catalases in
RT pumpkin.;
RL Plant Mol. Biol. 33:141-155(1997).
CC - FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC - SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC - CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC - COFACTOR: HEME GROUP (BY SIMILARITY).
CC - SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Glyoxysomal.
CC - TISSUE SPECIFICITY: HIGH EXPRESSION IN SEEDS AND EARLY SEEDLINGS.
CC - SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: D55645; BAA09506.1; -
DR HSP: P00432; 4BLG.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR ProDom: PD000510; Catalase; 1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; 1.
KW Oxidoreductase; peroxidase; iron; Heme; Hydrogen peroxide;
KW Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 492 AA; 57070 MW; B02C649F767FAC20 CRC64;
Query Match 47.8%; Score 75; DB 1; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.001;
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 9 TTAAGAPVWDDNNVITAGPRG 29
II IIIII::I::I::IIII
Db 18 TTNAGAPVNNSSMTVGPRG 38
RESULT 12
CAT2_ARATH STANDARD; PRT; 492 AA.
ID CAT2_ARATH

```

```

AC P25819; O49615;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase 2 (EC 1.11.1.6).
GN CAT2 OR CAT OR AT4G35090 OR M4E13.140.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Chevallier C., Yamaguchi J., McCourt P.;
RT "Nucleotide sequence of a cDNA for catalase from Arabidopsis
RT thaliana.";
RL Plant Physiol. 99:1726-1728(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG ERECTA;
RA Zentgraf U., Zinkernagel I.;
RT "A gene encoding a catalase isoform from Arabidopsis thaliana.";
RL (In) Plant Gene Register PGR96-005.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
RA Harris B., Ansorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidtheini T.,
RA Reichert B., Portetle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Honeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
RA Van der Schueren J., Grymonprez B., Chuang Y.-J., Vandenbussche F.,
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Deoor E.,
RA Weitzenecker T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
RA Bernheiser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
RA De Keyser A., Buyshaert C., Gielen J., Villarroel R., De Clercq R.,
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIlay K., Mayes R.,
RA Pettitt A., Rajandream M.-A., Lyne M., Benes V., Rechmann S.,
RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argiriou A., Vitale D., Liguori R., Piravandi E.,
RA Massenot O., Quigley F., Clabaud G., Muendlein A., Feilber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedior F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Baigues M., Terol J., Torres A.,
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Willson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
RA Steneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Fulton L., Mardis E., Dente M., Pepin K., Hillier L.,
RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Marra M., Martienssen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";

```

```

RL Nature 402:769-777 (1999).
CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER AND HETEROTETRAMER. AT LEAST SIX OR SEVEN
CC ISOZYMES ARE PRODUCED FROM A MIXTURE OF 3 GENE PRODUCTS.
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal.
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64271; CAA45564.1; -.
DR EMBL; X94447; CAA64220.1; -.
DR EMBL; AL022023; CAA17773.1; -.
DR EMBL; AL161586; CAB80226.1; -.
DR PIR; SI8972; SI8972.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT BINDING 138 138 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT VARIANT 421 421 I -> V (IN STRAIN CV. LANDSBERG ERECTA).
FT CONFLICT 39 39 P -> L (IN REF. 1).
FT CONFLICT 109 109 E -> A (IN REF. 1).
FT CONFLICT 154 154 M -> I (IN REF. 1).
FT CONFLICT 243 243 V -> L (IN REF. 1).
SQ SEQUENCE 492 AA; 56931 MW; AB622230561FD79B CRC64;

Query Match 47.8%; Score 75; DB 1; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.001;
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 9 TTAGGAPVWDDNNVITAGPRG 29
|| |||||::||:| |||||
Db 18 TTNSGAPVWNNSSMTVGPRG 38

RESULT 13
CAT1_HORVU
ID CAT1_HORVU STANDARD; PRT; 492 AA.
AC P55307; Q43761;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Catalase isozyme 1 (EC 1.11.1.6).
GN CAT1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MOREX;
RX MEDLINE=96145511; PubMed=8555444;
RA Skadsen R.W., Schulze-Lefert P., Herbst J.M.;
RT "Molecular cloning, characterization and expression analysis of two
RT catalase isozyme genes in barley.";
```

```

RL Plant Mol. Biol. 29:1005-1014(1995).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- TISSUE SPECIFICITY: IN WHOLE ENDOSPERMS (ALEURONES PLUS STARCHY
CC ENDOSPERM), IN ISOLATED ALEURONES AND IN DEVELOPING SEEDS.
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U20777; AAA96947.1; -.
DR EMBL; U16132; AAA62306.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT BINDING 138 138 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT BINDING 348 348
SQ SEQUENCE 492 AA; 56586 MW; 5917P285FD75D725 CRC64;

Query Match 47.1%; Score 74; DB 1; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.0015;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAGGAPVWDDNNVITAGPRG 29
|| |||||::||:| |||||
Db 18 TTNSGAPVWNNNALTGVGHRG 38

RESULT 14
CAT2_WHEAT
ID CAT2_WHEAT STANDARD; PRT; 492 AA.
AC P55313;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN CAT2.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Song Z., Zhu Y., Hui Liu G.J.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
```



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X94352; CAA64077.1; -  
HSPF: P00432; 4BLC.  
InterPro: IPR002226; Catalase.  
Pfam: PF00199; catalase; 1.  
PRINTS: PR00067; CATALASE.  
ProDom: PD000510; Catalase; 1.  
PROSITE: PS00437; CATALASE\_1; 1.  
PROSITE: PS00436; CATALASE\_2; 1.  
Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
Peroxisome; Glyoxysome.  
ACT\_SITE 65 65  
BY SIMILARITY.  
FT ACT\_SITE 138 138  
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 492 AA; 56480 MW; CFAED23345ED3C94 CRC64;

Query Match 47.1%; Score 74; DB 1; Length 492;  
Best Local Similarity 61.9%; Pred. No. 0.0015;  
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAAGAPVWDDNNVITAGPRG 29  
II IIIII::II::IIII  
Db 18 TTNSGAPVWNNNALTGVHGR 38

RESULT 15  
CATA\_DESVN  
ID CATA\_DESVN STANDARD; PRT; 479 AA.  
AC Q92N99;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Catalase (EC 1.11.1.6).  
GN KATA OR KAT.  
OS Desulfovibrio vulgaris (strain Miyazaki).  
OC Bacteria; Proteobacteria; delta subdivision; Desulfotomaculum.  
OX NCBI\_TaxID=883;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kitamura M., Kojima S., Akutsu H., Kumagai I., Nakaya T.;  
RT "Catalase from strictly anaerobic bacteria, Desulfovibrio vulgaris (Miyazaki F).";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.  
CC -1- CATALYTIC ACTIVITY: 2 H<sub>2</sub>O<sub>2</sub> = O<sub>2</sub> + 2 H<sub>2</sub>O.  
CC -1- COFACTOR: HEME GROUP.  
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AB020341; BAA34670.1; -  
HSPF: P42321; 2CAE.  
InterPro: IPR002226; Catalase.  
Pfam: PF00199; catalase; 1.  
PRINTS: PR00067; CATALASE.  
ProDom: PD000510; Catalase; 1.  
PROSITE: PS00437; CATALASE\_1; 1.  
PROSITE: PS00436; CATALASE\_2; 1.

Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.  
FT ACT\_SITE 54 54 BY SIMILARITY.  
FT ACT\_SITE 127 127 BY SIMILARITY.  
FT BINDING 337 337 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 479 AA; 54549 MW; 4CD97480AB6FD475 CRC64;

Query Match 46.5%; Score 73; DB 1; Length 479;  
Best Local Similarity 65.2%; Pred. No. 0.002;  
Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 7 KOTAFGAPVWDDNNVITAGPRG 29  
II IIIII::II::IIII  
Db 5 KLTTNAGAPVDPNQNTAGPRG 27

RESULT 16  
CATA\_PHAUU  
ID CATA\_PHAUU STANDARD; PRT; 492 AA.  
AC P32290;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Catalase (EC 1.11.1.6).  
OS Phaseolus aureus (mung bean) (Vigna radiata).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.  
OX NCBI\_TaxID=3916;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94151449; PubMed=8108520;  
RA Mori H., Imaseki H.;  
RT "cDNA for catalase from etiolated mung bean (Vigna radiata) Plant Physiol. 102:691-692(1993).  
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.  
CC -1- CATALYTIC ACTIVITY: 2 H<sub>2</sub>O<sub>2</sub> = O<sub>2</sub> + 2 H<sub>2</sub>O.  
CC -1- COFACTOR: HEME GROUP.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: D13557; BAA02755.1; -  
HSPF: P00432; 4BLC.  
InterPro: IPR002226; Catalase.  
Pfam: PF00199; catalase; 1.  
PRINTS: PR00067; CATALASE.  
ProDom: PD000510; Catalase; 1.  
PROSITE: PS00437; CATALASE\_1; 1.  
PROSITE: PS00436; CATALASE\_2; 1.  
Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
Peroxisome; Glyoxysome.  
ACT\_SITE 65 65 BY SIMILARITY.  
FT ACT\_SITE 138 138 BY SIMILARITY.  
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 492 AA; 56844 MW; 933E604D261ICE85 CRC64;

Query Match 46.5%; Score 73; DB 1; Length 492;  
Best Local Similarity 61.9%; Pred. No. 0.0021;  
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAAGAPVWDDNNVITAGPRG 29

```
Db 18 TTNSGAPVWNNNSLTGTRG 38
|| |||||:|:| |
|| |||||:|:| |

RESULT 17
CATA_METBA
ID CATA_METBA STANDARD; PRT; 505 AA.
AC 093662;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN KAT.
OS Methanosarcina barkeri.
OC Archaea; Euryarchaeota; Methanosarcinales; Methanosarcinaceae;
OC Methanosarcina.
OX NCBI_TaxID=2208;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FUSARO / DSM 804;
RA Shima S., Netrusov A., Sordel M., Wicke M., Hartmann G.C.,
RA Thauer R.K.;
RT "Purification, characterization and primary structure of a
RT monofunctional catalase from Methanosarcina barkeri.";
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ005939; CAA06774.1; -.
CC HSP; P04040; 100W
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
KW ACT_SITE 58 58 BY SIMILARITY.
FT ACT_SITE 131 131 BY SIMILARITY.
FT BINDING 341 341 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 505 AA; 57065 MW; 2A27C4BEC47BE854 CRC64;

Query Match 45.9%; Score 72; DB 1; Length 505;
Best Local Similarity 66.7%; Pred. No. 0.0031;
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29
|| |||||:|:| |
|| |||||:|:| |

Db 11 TTGFGIPVGDQNSLTAGNRG 31

RESULT 18
CATA_BACSU
ID CATA_BACSU STANDARD; PRT; 482 AA.
AC P26901; P77838;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vegetative catalase (EC 1.11.1.6).
GN KATA OR KAT OR KAT-19.
OC Bacillus subtilis.
OS
```

```
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / YB886;
RX MEDLINE=92097949; PubMed=1756979;
RA Bol D.K., Yasbin R.E.;
RT "The isolation, cloning and identification of a vegetative catalase
RT gene from Bacillus subtilis.";
RL Gene 109:31-37(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97346037; PubMed=9202460;
RA Cummings N.J., Connerton I.F.;
RT "The Bacillus subtilis 168 chromosome from ssPE to kata.";
RL Microbiology 143:1855-1859(1997).
RN [3]
RP SEQUENCE OF 1-18.
RC STRAIN=168 / YB886;
RX MEDLINE=94236234; PubMed=8180695;
RA Hartford O.M., Dowds B.C.A.;
RT "Isolation and characterization of a hydrogen peroxide resistant
RT mutant of Bacillus subtilis.";
RL Microbiology 140:297-304(1994).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- DEVELOPMENTAL STAGE: ELEVATED LEVELS OF EXPRESSION DURING ENTRY
CC INTO THE STATIONARY PHASE OF THE GROWTH CYCLE.
CC -1- INDUCTION: BY HYDROGEN PEROXIDE.
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M80796; AAA22402.1; -.
CC EMBL; Z82044; CAB04807.1; -.
CC EMBL; Z99108; CAB12710.1; -.
CC PIR; JH0532; JH0532.
CC HSP; P00432; 4BLC.
CC Subtilisin; BG10849; kata.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Complete proteome.
FT INIT_MET 0 0
FT ACT_SITE 53 53 BY SIMILARITY.
FT ACT_SITE 126 126 BY SIMILARITY.
FT BINDING 336 336 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 205 205 G -> P (IN REF. 1).
FT CONFLICT 372 372 G -> D (IN REF. 1).
SQ SEQUENCE 482 AA; 54602 MW; 749F7C88CEA2E047 CRC64;

Query Match 45.2%; Score 71; DB 1; Length 482;
Best Local Similarity 60.9%; Pred. No. 0.0041;
Matches 14; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 7 KQTAFGAPVDDNNVITAGPRG 29
| |:::| | | | | | | | | | | | | | | |
```

```
Db 4 KLFTSWGAPVGDNQNSMTAGSRG 26
RESULT 19
CATA_RHIME STANDARD; PRT; 494 AA.
AC P95631;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase A (EC 1.11.1.6).
GN KATA OR R00764 OR SMC00819.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=97113440; PubMed=8955300;
RA Herouart D., Sigaud S., Moreau S., Frendo P., Touati D., Puppo A.;
RT "Cloning and characterization of the kta gene of Rhizobium meliloti
encoding a hydrogen peroxide-inducible catalase.";
RL J. Bacteriol. 178:6802-6809(1996).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puchler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
Sinorhizobium meliloti strain 1021.";
RC Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U59271; AAC44649.1; -.
DR EMBL; AL591784; CAC45336.1; -.
DR HSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; FALSE_NEG.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Periplasmic; Complete proteome.
FT ACT_SITE 55 55 BY SIMILARITY.
FT ACT_SITE 127 127 BY SIMILARITY.
FT BINDING 337 337 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 1 19 MTDPTTTTAGAPVDNQ -> MPPPLPERHLLICACVGT
AGCGSNPPGSSSPFVPAEPTAEPKCGVSPRSR
AGFSPRIRALISSPVSVSYSSKPSARAT (IN REF. 1).
SQ SEQUENCE 494 AA; 56008 MW; 39FA86606D2D83D7 CRC64;

Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAGAPVWDNNVNITAGPRG 29
|| ||||| : |||||
Db 8 TTAGAPVDPNQNLSITAGPRG 28

RESULT 20
CATA_MAIZE STANDARD; PRT; 492 AA.
ID CATA_MAIZE
AC P18122;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catalase isozyme 1 (EC 1.11.1.6).
DE CATA.
GN Zea mays (Maize).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. RG-67; TISSUE=Scutellum;
RX MEDLINE=89051000; PubMed=2461221;
RA Redinbaugh M.G., Wadsworth G.J., Scandalios J.G.;
RT "Characterization of catalase transcripts and their differential
expression in maize.";
RL Biochim. Biophys. Acta 951:104-116(1988).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. W64A; TISSUE=Leaf;
RX MEDLINE=94035156; PubMed=8220459;
RA Guan L., Scandalios J.G.;
RT "Characterization of the catalase antioxidant defense gene Catl of
maize, and its developmentally regulated expression in transgenic
tobacco.";
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- TISSUE SPECIFICITY: SCUTELLA, MILKY ENDOSPERM OF IMMATURE
KERNELS, LEAVES AND EPICOTYLS.
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
DR EMBL; X12538; CAA31056.1; -.
DR EMBL; X60135; CAA42720.1; -.
DR PIR; JA0090; JA0090.
DR HSP; P21179; 1CF9.
DR MaizeDB; 13855; -.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT VARIANT 157 157 A -> V (IN STRAIN W64A).
```

FT VARIANT 211 211 S -> T (IN STRAIN W64A).  
 FT VARIANT 329 329 P -> I (IN STRAIN W64A).  
 FT VARIANT 483 483 P -> A (IN STRAIN W64A).  
 FT CONFLICT 332 332 A -> G (IN REF. 1).  
 FT CONFLICT 415 415 R -> G (IN REF. 1).  
 FT CONFLICT 456 456 H -> D (IN REF. 1).  
 SQ SEQUENCE 492 AA; 56877 MW; CE10C93BEC1D9529 CRC64;

Query Match 44.6%; Score 70; DB 1; Length 492;  
 Best Local Similarity 57.1%; Pred. No. 0.0059;  
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
 || |||||::: :| | ||  
 Db 18 TTNSGAPVWNNNSALTGVQRG 38

RESULT 21  
 CATB\_ORYSA STANDARD; PRT; 492 AA.  
 AC P55309;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Catalase isozyme B (EC 1.11.1.6) (CAT-B).  
 GN CATB.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RX MEDLINE=94336766; PubMed=80588828;  
 RA Morita S., Tasaka M., Fujisawa H., Ushimaru T., Tsuji H.;  
 RT "A cDNA clone encoding a rice catalase isozyme.";  
 RL Plant Physiol. 105:1015-1016(1994).  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE.  
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -!- COFACTOR: HEME GROUP.  
 CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D26484; BAA05494.1; -.  
 CC HSSP; P21179; ICF9.  
 CC InterPro; IPR002226; Catalase.  
 CC Pfam; PF00199; catalase; 1.  
 CC PRINTS; PR00067; CATALASE.  
 CC ProDom; PD000510; Catalase; 1.  
 CC PROSITE; PS00437; CATALASE\_1; 1.  
 CC PROSITE; PS00438; CATALASE\_2; 1.  
 CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Peroxisome; Glyoxysome; Multigene family.  
 FT ACT SITE 65 65 BY SIMILARITY.  
 FT ACT SITE 138 138 BY SIMILARITY.  
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 SQ SEQUENCE 492 AA; 56519 MW; 08963B05BA6C20AB CRC64;

Query Match 44.6%; Score 70; DB 1; Length 492;

Best Local Similarity 57.1%; Pred. No. 0.0059;  
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 9 TTAFGAPVDDNNVITAGPRG 29  
 || |||||::: :| | ||  
 Db 18 TTNSGAPVWNNNSALTGVQRG 38

RESULT 22  
 CATL\_ARATH STANDARD; PRT; 492 AA.  
 AC Q96528; O22529;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Catalase 1 (EC 1.11.1.6).  
 GN CATL.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANDSBERG ERECTA;  
 RX MEDLINE=96416445; PubMed=8819328;  
 RA Frugoli J.A., Zhong H.H., Nuccio M.L., McCourt P., McPeck M.A.,  
 RA Thomas T.L., McClung C.R.;  
 RT "Catalase is encoded by a multigene family in Arabidopsis thaliana  
 RT (L.) Heynh.";  
 RL Plant Physiol. 112:327-336(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. Columbia;  
 RX MEDLINE=98250696; PubMed=9584109;  
 RA Frugoli J.A., McPeck M.A., Thomas T.L., McClung C.R.;  
 RT "Intron loss and gain during evolution of the catalase gene family in  
 RT angiosperms.";  
 RL Genetics 149:355-365(1998).  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 CC PEROXIDE.  
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -!- COFACTOR: HEME GROUP.  
 CC -!- SUBUNIT: HOMOTETRAMER AND HETEROTETRAMER. AT LEAST SIX OR SEVEN  
 CC ISOZYMES ARE PRODUCED FROM A MIXTURE OF 3 GENE PRODUCTS.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U43340; AAB07026.1; -.  
 CC HSSP; AF021937; AAC17731.1; -.  
 CC HSSP; P00432; 4BLC.  
 CC InterPro; IPR002226; Catalase.  
 CC Pfam; PF00199; catalase; 1.  
 CC PRINTS; PR00067; CATALASE.  
 CC ProDom; PD000510; Catalase; 1.  
 CC PROSITE; PS00437; CATALASE\_1; 1.  
 CC PROSITE; PS00438; CATALASE\_2; 1.  
 CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Multigene family.  
 FT ACT SITE 65 65 BY SIMILARITY.  
 FT ACT SITE 138 138 BY SIMILARITY.  
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 103 103 F -> L (IN REF. 1).  
 FT CONFLICT 332 332 G -> F (IN REF. 1).

FT CONFLICT 339 340 KL -> NV (IN REF. 1).  
 FT CONFLICT 455 455 S -> W (IN REF. 1).  
 SQ SEQUENCE 492 AA; 56861 MW; 370F54FF7D757C9B CRC64;

Query Match 43.3%; Score 68; DB 1; Length 492;  
 Best Local Similarity 57.1%; Pred. No. 0.012;  
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFGAPVDDNNVITAGPRG 29  
 || |||||::: || || ||  
 Db 18 TTNSGAPVWNNSSLTGTRG 38

## RESULT 23

CAT1\_RICCO STANDARD; PRT; 492 AA.

ID CAT1\_RICCO  
 AC Q01297;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Catalase isozyme 1 (EC 1.11.1.6).  
 GN CAT1.  
 OS Ricinus communis (Castor bean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.  
 OX NCBI\_TaxID=3988;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hypocotyl;  
 RX MEDLINE=94325474; PubMed=8049373;  
 RA Suzuki M., Ario T., Hattori T., Nakamura K., Asahi T.;  
 RT "Isolation and characterization of two tightly linked catalase genes  
 from castor bean that are differentially regulated.";  
 RL Plant Mol. Biol. 25:507-516(1994).  
 RN [2]  
 RP SEQUENCE OF 404-492 FROM N.A.  
 RC STRAIN=CV HALE;  
 RX MEDLINE=91293125; PubMed=1712298;  
 RA Gonzalez E.;  
 RT "The C-terminal domain of plant catalases. Implications for a  
 glyoxysomal targeting sequence.";  
 RL Eur. J. Biochem. 193:211-215(1991).  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 PEROXIDE.  
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.  
 CC -!- COFACTOR: HEME GROUP.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (Potential).  
 CC -!- TISSUE SPECIFICITY: ABUNDANT IN ENDOSPERMS AND COTYLEDONS. ONLY IN  
 SMALL AMOUNT IN ROOT.  
 CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; D21161; BRA04697.1; -;  
 DR EMBL; X59694; CAA42215.1; -;  
 DR PIR; S16231; S16231.  
 DR PIR; S29333; S29333.  
 DR HSSP; P21179; 1CF9.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.

DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Peroxisome; Glyoxysome; Multigene family.  
 FT ACT\_SITE 65  
 FT ACT\_SITE 65 BY SIMILARITY.  
 FT BINDING 138 138 BY SIMILARITY.  
 FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 439 439 A -> P (IN REF. 2).  
 FT CONFLICT 446 446 F -> S (IN REF. 2).  
 FT CONFLICT 454 454 L -> F (IN REF. 2).  
 FT CONFLICT 456 456 D -> E (IN REF. 2).  
 FT CONFLICT 461 461 H -> Q (IN REF. 2).  
 SQ SEQUENCE 492 AA; 56464 MW; B5E28A425088BF63 CRC64;

Query Match 43.3%; Score 68; DB 1; Length 492;  
 Best Local Similarity 57.1%; Pred. No. 0.012;  
 Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFGAPVDDNNVITAGPRG 29  
 || |||||::: || || ||  
 Db 18 TTNSGAPVWNNSSLTGTRG 38

## RESULT 24

CAT2\_GOSHI STANDARD; PRT; 492 AA.

ID CAT2\_GOSHI  
 AC P30567;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 01-APR-1993 (Rel. 25, Last annotation update)  
 DE Catalase isozyme 2 (EC 1.11.1.6).  
 GN CAT2 OR SU2.  
 OS Gossypium hirsutum (Upland cotton).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Malvales; Malvaceae; Gossypium.  
 OX NCBI\_TaxID=3635;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV DELTAPINE 62; TISSUE=Cotyledon;  
 RX MEDLINE=91378551; PubMed=1898069;  
 RA Ni W., Trelease R.N.;  
 RT "Two genes encode the two subunits of cottonseed catalase.";  
 RL Arch. Biochem. Biophys. 289:237-243(1991).  
 CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
 SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
 PEROXIDE.  
 CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.  
 CC -!- COFACTOR: HEME GROUP.  
 CC -!- SUBUNIT: HOMOTETRAMER.  
 CC -!- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -!- MISCELLANEOUS: THERE ARE AT LEAST FIVE ISOZYMES OF CATALASE IN  
 COTTON.  
 CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; X56675; CAA39998.1; -;  
 DR PIR; S17493; S17493.  
 DR HSSP; P00432; 4BLC.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase; 1.  
 DR PRINTS; PR00067; CATALASE.  
 DR ProDom; PD000510; Catalase; 1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;

```

KW Peroxisome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 492 AA; 56935 MW; 589FD8DCA173AB12 CRC64;

Query Match 43.3%; Score 68; DB 1; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDNNVITAGPRG 29
|| |||||:::| | ||
Db 18 TTNSGAPVWNNSSLTVGARG 38

RESULT 25
CATA_SOLAP STANDARD; PRT; 492 AA.
AC O24339;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE Catalase (EC 1.11.1.6);
OS Soldanella alpina (Alpine snowbell).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericales; Primulaceae; Soldanella.
OX NCBI_TaxID=66308;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Schmidt M.;
RL Submitted M.;
RT "Isolation and characterization of a pea catalase cDNA.";
RT "Isolation and characterization of a pea catalase cDNA.";
RT Plant Mol. Biol. 17:1263-1265(1991).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBCELLULAR LOCATION: Peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X60169; CAA42736.1; -.
CC PIR; S18346; CSPM.
CC HSSP; P00432; 4BLC.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC ProDom; PD000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
CC Peroxisome.
CC ACT_SITE 65 65 BY SIMILARITY.
CC ACT_SITE 138 138 BY SIMILARITY.
CC BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
CC SEQUENCE 492 AA; 57344 MW; B5FC08E66B13B266 CRC64;

Query Match 43.3%; Score 68; DB 1; Length 494;
Best Local Similarity 57.1%; Pred. No. 0.012;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDNNVITAGPRG 29
|| |||||:::| | ||
Db 18 TTNSGAPVWNNSSLTVGSRG 38

RESULT 27
CATA_PSEPU STANDARD; PRT; 479 AA.
AC Q59714;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Catalase (EC 1.11.1.6).

```

```
GN KATA OR CATA.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CORVALLIS;
RX MEDLINE=98019091; PubMed=9358059;
RA Kim Y.C., Miller C.D., Anderson A.J.;
RT "Identification of adjacent genes encoding the major catalase and a
RT bacterioferritin from the plant-beneficial bacterium Pseudomonas
RT putida.";
RL Gene 199:219-224(1997).
CC -1- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- ENZYME REGULATION: ACTIVATED BY PEROXIDE.
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U63511; AAB88219.1; -
DR HSP; P42321; 2CAE.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PD00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide.
FT ACT_SITE 53 53 BY SIMILARITY.
FT ACT_SITE 126 126 BY SIMILARITY.
FT BINDING 336 336 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 479 AA; 53381 MW; EFE3CBDE67778571 CRC64;
-----
Query Match 42.7%; Score 67; DB 1; Length 479;
Best Local Similarity 66.7%; Pred. No. 0.016;
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 9 TTAAGAPVWDDNNVITAGPRG 29
||| |||| | : | |||||
DB 6 TTASGAPVADNQNSRSGPRG 26
-----
RESULT 28
CAT1_SOYBN
ID CAT1_SOYBN STANDARD; PRT; 492 AA.
AC P29756.
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase 1/2 (EC 1.11.1.6).
GN CAT1 AND CAT2.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV D8PL 415;
RX MEDLINE=92032793; PubMed=1932700;
RA Isin S.H., Allen R.D.;
RT "Isolation and characterization of a pea catalase cDNA.";
```

```
RL Plant Mol. Biol. 17:1263-1265(1991).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=CV. CORSOY 79;
RA Su H., Hardy K.A., Hermsmeider D., Baum T.J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -1- COFACTOR: HEME GROUP.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Z12021; CAA78056.1; -
DR EMBL; AF035252; AAB88169.1; -
DR EMBL; AF035253; AAB88170.1; -
DR PIR; S20999; CSSY.
DR HSP; P21179; ICF9.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PD00067; CATALASE.
DR PRODOM; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 492 AA; 56847 MW; CEGAEE8BEEA483C6 CRC64;
-----
Query Match 42.7%; Score 67; DB 1; Length 492;
Best Local Similarity 52.4%; Pred. No. 0.017;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 9 TTAAGAPVWDDNNVITAGPRG 29
||| |||| | : | |||||
DB 18 TTNSGAPIWNNNSLTGVRG 38
-----
RESULT 29
CAT3_NICPL
ID CAT3_NICPL STANDARD; PRT; 492 AA.
AC P49317;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Catalase isozyme 3 (EC 1.11.1.6).
GN CAT3.
OS Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4092;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Flower;
RX MEDLINE=95010681; PubMed=7925949;
RA Wilkens H., Villarroel R., van Montagu M., Inze D., van Camp W.;
RT "Molecular identification of catalases from Nicotiana plumbaginifolia
RT (L.).";
RL FEBS Lett. 352:79-83(1994).
```

```

CC CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN: SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- INDUCTION: BY 3-AMINOTRIAZOLE.
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z36977; CAA85426.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 492 AA; 57016 MW; 2A49AD89CC6FB4D5 CRC64;
SQ
Query Match 42.7%; Score 67; DB 1; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.017;
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 9 TTAFCGAPVWDDNNVITAGPRG 29
|| |||||:|:| |||
Db 18 TTNSGAPVWNNSSMTVGTGR 38

RESULT 30
CAT3_SOYBN
ID CAT3_SOYBN STANDARD; PRT; 492 AA.
AC O48560;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase 3 (EC 1.11.1.6).
GN CAT3.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Corsoy 79;
RA Su H., Hardy K.A., Hermsmeier D., Baum T.J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (By similarity).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF035254; AAB88171.1; -.
DR HSSP; P21179; 1CF9.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 492 AA; 56911 MW; 36265068CF9F6CF1 CRC64;
SQ
Query Match 42.7%; Score 67; DB 1; Length 492;
Best Local Similarity 52.4%; Pred. No. 0.017;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Oy 9 TTAFCGAPVWDDNNVITAGPRG 29
|| |||||:|:| |||
Db 18 TTNSGAPVWNNSSLTGVRG 38

RESULT 31
CAT4_SOYBN
ID CAT4_SOYBN STANDARD; PRT; 492 AA.
AC O48561;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase 4 (EC 1.11.1.6).
GN CAT4.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. Corsoy 79;
RA Su H., Hardy K.A., Hermsmeier D., Baum T.J.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (By similarity).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF035255; AAB88172.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.

```



DR PROSITE; PS00438; CATALASE.2; 1.  
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
KW Peroxisome; Glyoxysome; Multigene family.  
FT ACT\_SITE 65 BY SIMILARITY.  
FT ACT\_SITE 138 138 BY SIMILARITY.  
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 492 AA; 56737 MW; FEF3B4706A4FD669 CRC64;

Query Match 42.7%; Score 67; DB 1; Length 492;  
Best Local Similarity 52.4%; Pred. No. 0.017; Mismatches 5; Indels 0; Gaps 0;  
Matches 11; Conservative 5;

QY 9 TTAFGAPVWDDNNVITAGPRG 29  
IIIIII:::IIIIII  
DB 18 TTNSGAPIWNNSLTVGARG 38  
IIIIII:::IIIIII

RESULT 32  
CATA\_PIG STANDARD; PRT; 502 AA.  
AC 052839; 2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Catalase (EC 1.11.1.6).  
GN CAT.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-LANDRACE; TISSUE=Liver;  
RX MEDLINE=98136431; PubMed=9475954;  
RA Lin Z.-H., Wang Y.-F., Sarai A., Yasue H.;  
RT "Swine catalase deduced from cDNA and localization of the catalase  
gene on swine chromosome 2p16-p15.";  
RL Biochem. Genet. 35:297-302(1997).  
CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
CC PEROXIDE.  
CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) - O(2) + 2 H(2)O.  
CC -1- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
DR EMBL; D89812; BAA25301.1; -.  
DR HSPSP; P00432; 4BLC.  
DR InterPro; IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR PRODOM; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE.1; 1.  
DR PROSITE; PS00438; CATALASE.2; 1.  
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide; NADP;  
KW Peroxisome.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT ACT\_SITE 74 74 BY SIMILARITY.  
FT ACT\_SITE 147 147 BY SIMILARITY.  
FT BINDING 357 357 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 502 AA; 57182 MW; F40066B3C424212 CRC64;

Query Match 42.0%; Score 66; DB 1; Length 502;  
Best Local Similarity 57.1%; Pred. No. 0.024; Mismatches 3; Indels 0; Gaps 0;  
Matches 12; Conservative 3;

QY 9 TTAFGAPVWDDNNVITAGPRG 29  
IIIIII:::IIIIII  
DB 27 TTSGNPGDKLNLITAGPRG 47  
IIIIII:::IIIIII

RESULT 33  
CATA\_HUMAN STANDARD; PRT; 526 AA.  
AC P04040;  
DT 01-NOV-1986 (Rel. 03, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Catalase (EC 1.11.1.6).  
GN CAT.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=86286565; PubMed=3755526;  
RA Bell G. I., Najarian R. C., Mullenbach G. T., Hallemwell R. A.;  
RT "cDNA sequence coding for human kidney catalase.";  
RL Nucleic Acids Res. 14:5561-5562(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86286546; PubMed=3755525;  
RA Quan F., Korneluk R. G., Tropak M. B., Gravel R. A.;  
RT "Isolation and characterization of the human catalase gene.";  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 76-526 FROM N.A.  
RC TISSUE=Fibroblast;  
RX MEDLINE=85054813; PubMed=6548744;  
RA Korneluk R. G., Quan F., Lewis W. H., Guise K. S., Willard H. F.,  
RA Holmes M. T., Gravel R. A.;  
RT "Isolation of human fibroblast catalase cDNA clones. Sequence of  
RT clones derived from spliced and unspliced mRNA.";  
RL J. Biol. Chem. 259:13819-13823(1984).  
RN [5]  
RP SEQUENCE OF 1-21 FROM N.A.  
RX MEDLINE=94110338; PubMed=8282800;  
RA Yoo J.-H., Erzurum S. C., Hay J. G., Lemarchand P., Crystal R. G.;  
RT "Vulnerability of the human airway epithelium to hyperoxia.  
RT Constitutive expression of the catalase gene in human bronchial  
RT epithelial cells despite oxidant stress.";  
RL J. Clin. Invest. 93:297-302(1994).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.75 ANGSTROMS).  
RX MEDLINE=2013173; PubMed=10666617;  
RA Ko T. P., Safo M. K., Musayev F. N., Di Salvo M. L., Wang C., Wu S. H.,  
RA Abraham D. J.;  
RT "Structure of human erythrocyte catalase.";  
RL Acta Crystallogr. D 56:241-245(2000).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS).  
RX MEDLINE=20124014; PubMed=10656833;  
RA Putnam C. D., Arvai A. S., Bourne Y., Tainer J. A.;  
RT "Active and inhibited human catalase structures: ligand and NADPH  
RT binding and catalytic mechanism.";  
RL J. Mol. Biol. 296:295-309(2000).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
RX MEDLINE=21065390; PubMed=11134921;



DE Catalase (EC 1.11.1.6).  
GN KATA OR HKTE OR HI0928.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OX Haemophilus.  
RN NCBI\_TaxID=727;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=94245618; PubMed=8188593;  
RA Bishai W.R., Smith H.O., Barcak G.J.;  
RT "A peroxide/ascorbate-inducible catalase from Haemophilus influenzae  
is homologous to the Escherichia coli katE gene product.";  
RL J. Bacteriol. 176:2914-2921(1994).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES  
TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.  
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
CC -!- COFACTOR: HEME GROUP.  
CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
CC -!- INDUCTION: BY HYDROGEN PEROXIDE.  
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U02682; AAA20441.1; -;  
DR EMBL; U32774; AAC22587.1; -;  
DR HSSP; P42321; 2CAB.  
DR TIGR; HI0928; -;  
DR InterPro: IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR ProDom; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; 1.  
DR PROSITE; PS00438; CATALASE\_2; 1.  
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
KW Complete proteome.  
FT ACT\_SITE 63 63 BY SIMILARITY.  
FT ACT\_SITE 136 136 BY SIMILARITY.  
FT BINDING 346 346 PROXIMAL HEME LIGAND (BY SIMILARITY).  
FT CONFLICT 357 357 N -> C (IN REF. 1).  
FT CONFLICT 384 384 N -> T (IN REF. 1).  
FT CONFLICT 413 413 N -> S (IN REF. 1).  
FT CONFLICT 423 423 N -> T (IN REF. 1).  
SQ SEQUENCE 508 AA; 57680 MW; 4FAFC16DB44043B0 CRC64;

Query Match 41.4%; Score 65; DB 1; Length 508;  
Best Local Similarity 61.9%; Pred. No. 0.035;  
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFGAPVVDNNVITAGPRG 29  
| | | | | : | : | | | | |  
DB 16 TNGNGAPVADNQNSLTAGPRG 36  
| | | | | : | : | | | | |  
RESULT 36  
CATA\_CANFA  
ID CATA\_CANFA STANDARD; PRT; 526 AA.  
AC Q97492;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Catalase (EC 1.11.1.6).  
GN CAT.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BEAGLE; TISSUE=Liver;  
RA Nakamura K.;  
RT "cDNA sequence and deduced amino acid sequence of dog catalase.";  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND  
SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN  
PEROXIDE.  
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
CC -!- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).  
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB012918; BAA36420.1; -;  
DR HSSP; P00432; 4BLC.  
DR InterPro: IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR ProDom; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; 1.  
DR PROSITE; PS00438; CATALASE\_2; 1.  
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide; NADP;  
KW Peroxisome.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT ACT\_SITE 74 74 BY SIMILARITY.  
FT ACT\_SITE 147 147 BY SIMILARITY.  
FT BINDING 357 357 PROXIMAL HEME LIGAND (BY SIMILARITY).  
SQ SEQUENCE 526 AA; 59666 MW; CB00F7F88FE2D972 CRC64;  
Query Match 41.4%; Score 65; DB 1; Length 526;  
Best Local Similarity 61.9%; Pred. No. 0.037;  
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFGAPVVDNNVITAGPRG 29  
| | | | | : | : | | | | |  
DB 27 TTGGGNPIGDKLNVTAGPRG 47  
| | | | | : | : | | | | |  
RESULT 37  
CATA\_CANFO  
ID CATA\_CANFO STANDARD; PRT; 526 AA.  
AC Q64405; Q9Q251;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN CAT.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Pinteric M., Baumgart E., Bulitta C., Fahimi D., Voelkl A.;
RT "Molecular characterization of guinea pig catalase.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-137 FROM N.A.
RC STRAIN=HARTLEY; TISSUE=Lung;
RX MEDLINE=96180320; PubMed=8597602;
RA Yuan H.T., Bingle C.D., Kelly F.J.;
RT "Differential patterns of antioxidant enzyme mRNA expression in
RT guinea pig lung and liver during development.";
RL Biochim. Biophys. Acta 1305:163-171(1996).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DE EMBL; AJ005111; CAB57222.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide; NADP;
KW Peroxisome.
FT INIT_MET 0 0 BY SIMILARITY.
FT ACT_SITE 74 74 BY SIMILARITY.
FT ACT_SITE 147 147 BY SIMILARITY.
FT BINDING 357 357 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 8 8 MISSING (IN REF. 2).
FT CONFLICT 82 82 A -> G (IN REF. 2).
SQ SEQUENCE 526 AA; 59802 MW; 7B58FDD76255E276 CRC64;

Query Match 41.4%; Score 65; DB 1; Length 526;
Best Local Similarity 61.9%; Pred. No. 0.037;
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 9 TTAAGPVDNNVITAGPRG 29
||| ||| | ||| |||
Db 27 TTAGNPVGDKNINVTGPRG 47

RESULT 38
CATL_WHEAT
ID CATL_WHEAT STANDARD; PRT; 492 AA.
AC Q43206;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

```

DE Catalase 1 (EC 1.11.1.6).
GN CAT1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Saruyama H., Matsumura T.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) -> O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal and glyoxysomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DE EMBL; D86327; BAAL3068.1; -.
DR HSSP; P00432; 4BLC.
DR InterPro; IPR002226; Catalase.
DR Pfam; PF00199; catalase; 1.
DR PRINTS; PR00067; CATALASE.
DR ProDom; PD000510; Catalase; 1.
DR PROSITE; PS00437; CATALASE_1; 1.
DR PROSITE; PS00438; CATALASE_2; 1.
KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; Glyoxysome; Multigene family.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
SQ SEQUENCE 492 AA; 56808 MW; 52899174D5BA2EAB CRC64;

Query Match 40.8%; Score 64; DB 1; Length 492;
Best Local Similarity 52.4%; Pred. No. 0.048;
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAAGPVDNNVITAGPRG 29
:| |||||::| |||
Db 18 STNSGAPVWNNDSLTVGSRG 38

RESULT 39
CATA_HELAN
ID CATA_HELAN STANDARD; PRT; 492 AA.
AC P45739;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Catalase (EC 1.11.1.6).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
OC Heliantheae; Helianthus.
OX NCBI_TaxID=4232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SPANNERS ALLZWECK; TISSUE=Cotyledon;
RX MEDLINE=95101741; PubMed=7803505;
RA Kleff S., Trelease R.N., Eising R.;
RT "Nucleotide and deduced amino acid sequence of a putative higher

```

```

RT molecular weight precursor for catalase in sunflower cotyledons."
RL Biochim. Biophys. Acta 1224:463-466(1994).
CC -!- FUNCTION: DECOMPOSES HYDROGEN PEROXIDE IN WATER AND OXYGEN; SERVES
CC TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.
CC -!- COFACTOR: HEME GROUP.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L28740; AAA69866.1; -.
CC HSSP; P00432; 4BLC.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; Catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC PROSITE; PS000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome.
FT ACT_SITE 65 65 BY SIMILARITY.
FT ACT_SITE 138 138 BY SIMILARITY.
FT BINDING 348 348 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT SEQUENCE 492 AA; 56755 MW; 4A5C596506F44961 CRC64;
SQ
Query Match 40.8%; Score 64; DB 1; Length 492;
Best Local Similarity 57.1%; Pred. No. 0.048; 5; Indels 0; Gaps 0;
Matches 12; Conservative 4; Mismatches 5;
Qy 9 TTAFGAPVWDDNNVITAGPRG 29
|| |||||:|:|:|
Db 18 TTNSGAPVYNNNSITVGSRG 38

RESULT 40
CATA_BRARE STANDARD; PRT; 526 AA.
AC O9PT92; Q9I8V5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Catalase (EC 1.11.1.6).
GN CAT.
OS Brachydanio rerio (Zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Ken C.F., Lin C.T., Wu J.L., Shaw J.F.;
RT "Molecular cloning of a cDNA coding for catalase from zebrafish (Danio rerio).";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gerhard G.S., Kauffman E.J., Grundy M.A.;
RT "Molecular cloning and sequence analysis of the Danio rerio catalase gene.";
RL Comp. Biochem. Physiol. 127:447-457(2000).
CC -!- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND
CC SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN
CC PEROXIDE.
CC -!- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.

```

```

CC -!- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Peroxisomal (By similarity).
CC -!- SIMILARITY: BELONGS TO THE CATALASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ007505; CAB64949.1; -.
CC EMBL; AF170069; AAF89686.1; -.
CC HSSP; P00432; 4BLC.
CC ZFIN; ZDB-GENE-000210-20; cat.
CC InterPro; IPR002226; Catalase.
CC Pfam; PF00199; catalase; 1.
CC PRINTS; PR00067; CATALASE.
CC PROSITE; PS000510; Catalase; 1.
CC PROSITE; PS00437; CATALASE_1; 1.
CC PROSITE; PS00438; CATALASE_2; 1.
CC Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;
KW Peroxisome; NADP.
FT ACT_SITE 75 75 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
FT BINDING 358 358 PROXIMAL HEME LIGAND (BY SIMILARITY).
FT CONFLICT 110 110 V -> A (IN REF. 2).
FT CONFLICT 123 123 P -> S (IN REF. 2).
FT CONFLICT 152 152 T -> I (IN REF. 2).
FT CONFLICT 161 161 S -> F (IN REF. 2).
FT CONFLICT 350 352 MLQ -> NAA (IN REF. 2).
FT CONFLICT 478 478 M -> T (IN REF. 2).
SQ SEQUENCE 526 AA; 59654 MW; E1120D3796522785 CRC64;
Query Match 40.8%; Score 64; DB 1; Length 526;
Best Local Similarity 57.1%; Pred. No. 0.052;
Matches 12; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Qy 9 TTAFGAPVWDDNNVITAGPRG 29
|| |||:|:|:|
Db 28 TTGAGVPIGDKLNAMTAGPRG 48

```

Search completed: August 22, 2002, 07:45:39  
Job time: 158 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:44:04 ; Search time 51.87 Seconds  
(without alignments)  
62.100 Million cell updates/sec

Title: US-09-488-737-2

Perfect score: 157

Sequence: 1 MVNKKVQTTAFGAPVDDNNVITAGPRG 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A\_Geneseq\_032802.\*

1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	157	100.0	29	16	AA1988244
2	157	100.0	29	18	AA1983405
3	157	100.0	505	18	AA1986481
4	157	100.0	505	18	AA1986482
5	157	100.0	505	18	AA1986483
6	157	100.0	505	19	AA1989423
7	157	100.0	505	19	AA1989423
8	153	97.5	505	19	AA1989423
9	75	47.8	377	21	AA1982827
10	75	47.8	394	21	AA1982826
11	75	47.8	492	20	AA1981831
					H. pylori catalase
					N-terminal sequenc
					Helicobacter CP2 a
					Helicobacter CP2-p
					Helicobacter CP2-p
					H. pylori GPHO 358
					Helicobacter pylori
					Helicobacter pylori
					Arabidopsis thalia
					Arabidopsis thalia
					Clone of A. thalia

12	75	47.8	492	21	AAG43132	Arabidopsis thalia
13	75	47.8	512	21	AA1983903	Arabidopsis thalia
14	71	45.2	483	21	AA1983903	Bacillus thermogl
15	70	44.6	115	21	AA1983903	Zea mays protein f
16	68	43.3	492	21	AA1983903	Arabidopsis thalia
17	68	43.3	536	21	AA1983903	Arabidopsis thalia
18	67	42.7	15	19	AA1983903	H. pylori 63 kD ad
19	66	42.0	57	22	AA1983903	Peptide #3098 enco
20	66	42.0	57	22	AA1983903	Peptide #3121 enco
21	66	42.0	57	22	AA1983903	Protein #3043 enco
22	66	42.0	57	22	AA1983903	Human bone marrow
23	66	42.0	57	22	AA1983903	Human bone marrow
24	66	42.0	57	22	AA1983903	Peptide #3065 enco
25	66	42.0	57	22	AA1983903	Peptide #3154 enco
26	66	42.0	57	22	AA1983903	Peptide #3030 enco
27	66	42.0	483	20	AA1983903	Bacillus thermogl
28	66	42.0	483	21	AA1983903	Bacillus thermogl
29	66	42.0	509	22	AA1983903	V. rumoiensis S-1
30	66	42.0	527	9	AA1983903	Recombinant cata
31	66	42.0	527	19	AA1983903	Catalase. Homo sa
32	66	42.0	527	20	AA1983903	Human catalase seq
33	66	42.0	527	22	AA1983903	Human protein SEQ
34	66	42.0	527	22	AA1983903	Human polypeptide
35	66	42.0	558	22	AA1983903	Human polypeptide
36	66	42.0	565	22	AA1983903	Human protein SEQ
37	64	40.8	492	19	AA1983903	Wheat catalase pro
38	61	38.9	499	21	AA1983903	Rat mitogenic regu
39	61	38.9	563	21	AA1983903	Rat mitogenic regu
40	58	36.9	152	20	AA1983903	Human prostate tum
41	58	36.9	332	21	AA1983903	Human mitogenic re
42	58	36.9	483	20	AA1983903	Bacillus thermogl
43	58	36.9	564	21	AA1983903	Human mitogenic re
44	57	36.3	527	17	AA1983903	Amino acid sequenc
45	56	35.7	527	17	AA1983903	Mouse catalase. M
46	54	34.4	489	15	AA1983903	Partial sequence o
47	54	34.4	489	16	AA1983903	SABP encoded by PC
48	52.5	33.4	729	14	AA1983903	Catalase-R. Asper
49	52.5	33.4	730	14	AA1983903	Aspergillus niger
50	52	33.1	163	22	AA1983903	C glutamicum prote
51	51	32.5	377	22	AA1983903	Novel human diagno
52	51	32.5	600	22	AA1983903	Novel human diagno
53	50	31.8	489	17	AA1983903	Tobacco salicylic
54	50	31.8	1188	19	AA1983903	Amino acid sequenc
55	50	31.8	1216	17	AA1983903	Human tyrosine dep
56	49	31.2	127	22	AA1983903	Propionibacterium
57	49	31.2	676	20	AA1983903	Human 15S lipoxyme
58	49	31.2	676	22	AA1983903	Human 15S-lipoxyme
59	48	30.6	354	22	AA1983903	Murine Wnt-8b prot
60	48	30.6	415	22	AA1983903	Human Wnt3 (2Wnt3)
61	48	30.6	502	22	AA1983903	C glutamicum prote
62	48	30.6	506	22	AA1983903	Drosophila melanog
63	48	30.6	516	22	AA1983903	C. glutamicum SRT
64	48	30.6	4601	22	AA1983903	Drosophila melanog
65	47.5	30.3	1428	22	AA1983903	Sepp protein enco
66	47	29.9	118	21	AA1983903	Arabidopsis thalia
67	47	29.9	135	21	AA1983903	Arabidopsis thalia
68	47	29.9	187	21	AA1983903	Arabidopsis thalia
69	47	29.9	323	22	AA1983903	Novel human diagno
70	46.5	29.6	279	17	AA1983903	Proteinase K varia
71	46.5	29.6	753	22	AA1983903	Novel human diagno
72	46.5	29.6	1094	22	AA1983903	Propionibacterium
73	46	29.3	56	22	AA1983903	Human secreted pro
74	46	29.3	153	21	AA1983903	Human protein sequ
75	46	29.3	153	22	AA1983903	Canine hookworm ne
76	46	29.3	248	20	AA1983903	Propionibacterium
77	46	29.3	508	22	AA1983903	Propionibacterium
78	46	29.3	636	20	AA1983903	Rhodococcus corall
79	46	29.3	637	21	AA1983903	Rhodococcus sp. oh
80	46	29.3	750	22	AA1983903	Salmonella typhi c
81	46	29.3	1000	22	AA1983903	Amino acid sequenc
82	45.5	29.0	298	21	AA1983903	Arabidopsis thalia
83	45.5	29.0	298	21	AA1983903	Arabidopsis thalia
84	45.5	29.0	326	21	AA1983903	Arabidopsis thalia

85 45.5 29.0 445 22 AAB93714 Human protein sequ  
 86 45.5 29.0 672 21 AAG52460 Arabidopsis thalia  
 87 45.5 29.0 680 21 AAG52462 Arabidopsis thalia  
 88 45.5 29.0 721 21 AAG52459 Arabidopsis thalia  
 89 45.5 29.0 729 21 AAG45661 Arabidopsis thalia  
 90 45.5 29.0 751 22 AAM78945 Human protein SEQ  
 91 45.5 29.0 751 22 AAM39546 Human polypeptide  
 92 45.5 29.0 775 21 AAG52458 Arabidopsis thalia  
 93 45.5 29.0 783 21 AAG45660 Arabidopsis thalia  
 94 45.5 29.0 802 22 AAM79929 Human protein SEQ  
 95 45.5 29.0 802 22 AAM41332 Human polypeptide  
 96 45 28.7 187 22 ABB57901 Drosophila melanog  
 97 45 28.7 189 22 AAY72770 Candida albicans,  
 98 45 28.7 316 13 AAR22277 Human pancreatic G  
 99 45 28.7 316 16 AAR75111 Glycosyl-phosphati  
 100 45 28.7 417 22 AAB92908 Human protein sequ

## ALIGNMENTS

RESULT 1  
 AAR8244  
 ID AAR8244 standard; peptide; 29 AA.  
 XX  
 AC AAR8244;  
 XX  
 DT 17-MAY-1996 (first entry)  
 XX  
 DE H.pylori catalase N-terminus.  
 XX  
 KW Helicobacter pylori; catalase; immunisation; gastritis; peptic ulcer;  
 KW stomach cancer; mutant; antigen; urease; encapsulate; liposome.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN W09527506-A1.  
 XX  
 PD 19-OCT-1995.  
 XX  
 PF 27-MAR-1995; 95WO-FR00383.  
 XX  
 PR 08-APR-1994; 94FR-0004172.  
 XX  
 PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.  
 XX  
 PI Lissolo L;  
 XX  
 DR WPI; 1995-366230/47.  
 XX  
 PT Vaccines against Helicobacter pylori - contg. H. pylori catalase  
 PS Example 1; Page 4; 14pp; French.  
 CC The N-terminal sequence of the 54 kD monomeric subunit of the  
 CC Helicobacter pylori ATCC 43579 catalase. The complete catalase (a  
 CC tetrameric protein) can be used in compns. for immunising against  
 CC H.pylori, pref. those implicated in gastritis, peptic ulcers and  
 CC stomach cancer. The compns. contain inactive mutant forms of the  
 CC catalase and may also contain adjuvants e.g. alum, and other H.pylori  
 CC antigens e.g. urease. The compns. are pref. encapsulated in liposomes.  
 XX  
 SQ Sequence 29 AA;  
 Query Match 100.0%; Score 157; DB 16; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-17;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MVNKKVQKTTAFGAPVWDDNNVITAGPRG 29  
 Db 1 mvnkdvkgttafgapvwdnnvtagprg 29  
 |||||||||||||||||||||||||||||

RESULT 2  
 AAW33405  
 ID AAW33405 standard; peptide; 29 AA.  
 XX  
 AC AAW33405;  
 XX  
 DT 05-MAR-1998 (first entry)  
 XX  
 DE N-terminal sequence of 54 kD membrane protein of H. pylori.  
 XX  
 KW membrane protein; vaccine; immunisation; immunoassay;  
 KW diagnostic; affinity purification; gastritis; peptic ulcer;  
 KW gastric cancer.  
 XX  
 OS Helicobacter pylori.  
 XX  
 PN W09712909-A1.  
 XX  
 PD 10-APR-1997.  
 XX  
 PF 04-OCT-1996; 96WO-FR01552.  
 XX  
 PR 04-OCT-1995; 95FR-0011890.  
 XX  
 PA (INMR ) PASTEUR MERIEUX SERUMS & VACCINS.  
 XX  
 PI Lissolo L;  
 XX  
 DR WPI; 1997-226163/20.  
 XX  
 PT New Helicobacter pylori membrane proteins - and related  
 PT polypeptide(s) and antibodies, useful for active or passive  
 PT immunisation and diagnosis  
 XX  
 PS Example 6; Page 25; 48pp; French.  
 XX  
 CC New purified Helicobacter pylori proteins are obtained from a  
 CC membrane fraction and have apparent molecular weights (by  
 CC electrophoresis on= 10% polyacrylamide gel in presence of SDS) of 54,  
 CC 50, 32-35 and 30 kD respectively, of which the 54 kD protein does not  
 CC react with anti-catalase antiserum. The present sequence represents the  
 CC N-terminal of the 54 kD protein. The proteins and purified peptides  
 CC derived from them can be used (by active immunisation) to treat or  
 CC prevent H. pylori infections (gastritis, peptic ulcers, gastric cancer  
 CC etc.). Antibodies raised against the protein can be used for passive  
 CC immunisation. Derived peptides and antibodies are also useful as  
 CC reagents for diagnostic immunoassays, and antibodies are further useful  
 CC for affinity purification of the new proteins.  
 XX  
 SQ Sequence 29 AA;  
 Query Match 100.0%; Score 157; DB 18; Length 29;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-17;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MVNKKVQKTTAFGAPVWDDNNVITAGPRG 29  
 Db 1 mvnkdvkgttafgapvwdnnvtagprg 29  
 |||||||||||||||||||||||||||||

RESULT 3  
 AAW06481  
 ID AAW06481 standard; Protein; 505 AA.  
 XX  
 AC AAW06481;  
 XX  
 DT 24-FEB-1997 (first entry)  
 XX  
 DE Helicobacter CP2 antigen.  
 XX  
 KW CP2; antigen; vaccine; diagnosis; gastritis.



XX OS Helicobacter pylori strain ATCC 43504.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 130 /note= "residue 130 is Asp in CP2-PCR2"  
 FT Misc-difference 160 /note= "residue 160 is Pro in CP2-PCR2"  
 FT Misc-difference 218 /note= "residue 218 is Leu in CP2-PCR2"  
 FT Misc-difference 234 /note= "residue 234 is Ile in CP2-PCR2"  
 FT Misc-difference 237 /note= "residue 237 is His in CP2-PCR2"  
 FT Misc-difference 248 /note= "residue 248 is Asp in CP2-PCR2"  
 FT Misc-difference 262 /note= "residue 262 is Ile in CP2-PCR2"  
 FT Misc-difference 285 /note= "residue 285 is Cys in CP2-PCR2"  
 FT Misc-difference 287 /note= "residue 287 is Lys in CP2-PCR2"  
 FT Misc-difference 292 /note= "residue 292 is Thr in CP2-PCR2"  
 FT Misc-difference 316 /note= "residue 316 is Thr in CP2-PCR2"  
 FT Misc-difference 344 /note= "residue 344 is Cys in CP2-PCR1"  
 FT Misc-difference 358 /note= "residue 358 is Arg in CP2-PCR2"  
 FT Misc-difference 397 /note= "residue 397 is Thr in CP2-PCR2"  
 FT Misc-difference 449 /note= "residue 449 is Gly in CP2-PCR1 and CP2-PCR2"  
 XX EP745674-A2.  
 XX PD 04-DEC-1996.  
 XX PF 30-MAY-1996; 96EP-0108637.  
 XX PR 05-APR-1996; 96JP-0083512.  
 XX PR 02-JUN-1995; 95JP-0136564.  
 XX PA (WAKP ) WAKO PURE CHEM IND LTD.  
 XX PI Hirayasu K, Kawabata T, Sugiyama T, Tanaka T;  
 XX DR WPI; 1997-013697/02.  
 XX DR N-PSDB; AAT45041.  
 XX PT Helicobacter pylori CP2 antigen structural polypeptide(s) - also  
 XX PT corresponding DNA, useful as reagents for detecting H. pylori and in  
 XX PT the production of vaccines  
 XX PS Claim 1; Page 16-18; 29pp; English.  
 XX CC Helicobacter pylori ATCC 43504 CP2 antigen (AAW06481) is useful in  
 XX CC the specific and quantitative detection of H. pylori and in vaccine  
 XX CC prodn. CP2 antibody titre shows high correlation with the  
 XX CC pathology of gastritis. The CP2 amino acid sequence was deduced  
 XX CC from a genomic clone (AAT45042). Slightly different sequences,  
 XX CC CP2-PCR1 (AAW06482) and CP2-PCR2 (AAW06483), were deduced from clones  
 XX CC (AAT45042-43) obt'd. by PCR amplification. Isolation of the gene  
 XX CC sequences allows the mass prodn. of recombinant CP2 in transformed  
 XX CC host cells.  
 XX SQ Sequence 505 AA;  
 Query Match 100.0%; Score 157; DB 18; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-15;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNKDVKQTTAFGAPVWDNNVITAGPRG 29  
 Db 1 mvnkdvkqtafgapvwdnnvtagprg 29  
 RESULT 4  
 AAW06482  
 ID AAW06482 standard; Protein; 505 AA.  
 XX AC AAW06482;  
 XX DT 24-FEB-1997 (first entry)  
 XX DE Helicobacter CP2-PCR1 antigen.  
 XX KW CP2; antigen; vaccine; diagnosis; gastritis.  
 XX OS Helicobacter pylori strain ATCC 43504.  
 XX FH Key Location/Qualifiers  
 FT Misc-difference 130 /note= "residue 130 is Asp in CP2-PCR2"  
 FT Misc-difference 160 /note= "residue 160 is Pro in CP2-PCR2"  
 FT Misc-difference 218 /note= "residue 218 is Leu in CP2-PCR2"  
 FT Misc-difference 234 /note= "residue 234 is Ile in CP2-PCR2"  
 FT Misc-difference 237 /note= "residue 237 is His in CP2-PCR2"  
 FT Misc-difference 248 /note= "residue 248 is Asp in CP2-PCR2"  
 FT Misc-difference 262 /note= "residue 262 is Ile in CP2-PCR2"  
 FT Misc-difference 285 /note= "residue 285 is Cys in CP2-PCR2"  
 FT Misc-difference 287 /note= "residue 287 is Lys in CP2-PCR2"  
 FT Misc-difference 292 /note= "residue 292 is Thr in CP2-PCR2"  
 FT Misc-difference 316 /note= "residue 316 is Thr in CP2-PCR2"  
 FT Misc-difference 344 /note= "residue 344 is Arg in CP2-PCR2 and CP2"  
 FT Misc-difference 358 /note= "residue 358 is Arg in CP2-PCR2"  
 FT Misc-difference 397 /note= "residue 397 is Thr in CP2-PCR2"  
 FT Misc-difference 449 /note= "residue 449 is Arg in CP2"  
 XX EP745674-A2.  
 XX PD 04-DEC-1996.  
 XX PF 30-MAY-1996; 96EP-0108637.  
 XX PR 05-APR-1996; 96JP-0083512.  
 XX PR 02-JUN-1995; 95JP-0136564.  
 XX PA (WAKP ) WAKO PURE CHEM IND LTD.  
 XX PI Hirayasu K, Kawabata T, Sugiyama T, Tanaka T;  
 XX DR WPI; 1997-013697/02.  
 XX DR N-PSDB; AAT45042.  
 XX PT Helicobacter pylori CP2 antigen structural polypeptide(s) - also  
 XX PT corresponding DNA, useful as reagents for detecting H. pylori and in  
 XX PT the production of vaccines

XX PS Claim 1; Page 18-20; 29pp; English.  
 XX CC Helicobacter pylori ATCC 43504 CP2-PCR1 antigen (AAW06482) is useful  
 CC in the specific and quantitative detection of H. pylori and in  
 CC vaccine prodn. CP2 antibody titre shows high correlation with the  
 CC pathology of gastritis. The CP2-PCR1 amino acid sequence was  
 CC deduced from a DNA clone (AAT45042) obtd. by PCR amplification.  
 CC Slightly different sequences were deduced for CP2 (AAW06481) from  
 CC genomic DNA (AAT45041) and for CP2-PCR2 (AAW06483) from a PCR fragment  
 CC (AAT45043) derived from a clinical isolate. Isolation of the gene  
 CC sequences allows the mass prodn. of recombinant CP2 in transformed  
 CC host cells.  
 XX SQ Sequence 505 AA;

Query Match 100.0%; Score 157; DB 18; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNKKVKQTTAFGAPVWDDNNVITAGPRG 29  
 |||||  
 Db 1 mvnkdvkqttafgavpwwdnnvtagprg 29

RESULT 5  
 AAW06483  
 ID AAW06483 standard; Protein; 505 AA.  
 XX AC AAW06483;

XX DT 24-FEB-1997 (first entry)  
 XX DE Helicobacter CP2-PCR2 antigen.  
 XX KW CP2; antigen; vaccine; diagnosis; gastritis.  
 XX OS Helicobacter pylori.

XX FH Key Location/Qualifiers  
 FT Misc-difference 130 /note= "residue 130 is Asn in CP2 and CP2-PCR1"  
 FT Misc-difference 160 /note= "residue 160 is His in CP2 and CP2-PCR1"  
 FT Misc-difference 218 /note= "residue 218 is His in CP2 and CP2-PCR1"  
 FT Misc-difference 234 /note= "residue 234 is Val in CP2 and CP2-PCR1"  
 FT Misc-difference 237 /note= "residue 237 is Tyr in CP2 and CP2-PCR1"  
 FT Misc-difference 248 /note= "residue 248 is Asn in CP2 and CP2-PCR1"  
 FT Misc-difference 262 /note= "residue 262 is Val in CP2 and CP2-PCR1"  
 FT Misc-difference 285 /note= "residue 285 is Tyr in CP2 and CP2-PCR1"  
 FT Misc-difference 287 /note= "residue 287 is Gln in CP2 and CP2-PCR1"  
 FT Misc-difference 292 /note= "residue 292 is Met in CP2 and CP2-PCR1"  
 FT Misc-difference 316 /note= "residue 316 is Ser in CP2 and CP2-PCR1"  
 FT Misc-difference 344 /note= "residue 344 is Cys in CP2-PCR1"  
 FT Misc-difference 358 /note= "residue 358 is Lys in CP2 and CP2-PCR1"  
 FT Misc-difference 397 /note= "residue 397 is Ala in CP2 and CP2-PCR1"  
 FT Misc-difference 449 /note= "residue 449 is Arg in CP2"  
 XX PN EP745674-A2.

XX PD 04-DEC-1996.  
 XX PF 30-MAY-1996; 96EP-0108637.  
 XX PR 05-APR-1996; 96JP-0083512.  
 XX PR 02-JUN-1995; 95JP-0136564.  
 XX (WAKP ) WAKO PURE CHEM IND LTD.  
 XX HIRAYASU K, Kawabata T, Sugiyama T, Tanaka T;  
 DR WPI; 1997-013697/02.  
 DR N-PSDB; AAT45043.  
 XX Helicobacter pylori CP2 antigen structural polypeptide(s) - also  
 FT corresponding DNA, useful as reagents for detecting H. pylori and in  
 FT the production of vaccines  
 XX PS Claim 1; Page 21-23; 29pp; English.  
 XX CC Helicobacter pylori CP2-PCR2 antigen (AAW06483) is useful in the  
 CC specific and quantitative detection of H. pylori and in vaccine  
 CC prodn. CP2 antibody titre shows high correlation with the  
 CC pathology of gastritis. The CP2-PCR2 amino acid sequence was  
 CC deduced from a DNA clone (AAT45043) obtd. by PCR amplification of  
 CC DNA from H. pylori isolated from a patient with chronic gastritis.  
 CC Slightly different sequences were deduced for CP2 (AAW06481-82) from  
 CC H. pylori ATCC 43504. Isolation of these gene allows the mass  
 CC prodn. of recombinant CP2 in transformed host cells.  
 XX SQ Sequence 505 AA;  
 Query Match 100.0%; Score 157; DB 18; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MVNKKVKQTTAFGAPVWDDNNVITAGPRG 29  
 |||||  
 Db 1 mvnkdvkqttafgavpwwdnnvtagprg 29  
 RESULT 6  
 AAW98423  
 ID AAW98423 standard; Protein; 505 AA.  
 XX AC AAW98423;  
 XX DT 31-MAR-1999 (first entry)  
 XX DE H. pylori GHPO 358 protein.  
 XX KW GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;  
 XX KW peptic ulcer disease.  
 XX OS Helicobacter pylori.  
 XX PN WO9843478-A1.  
 XX PD 08-OCT-1998.  
 XX PF 01-APR-1998; 98WO-US06371.  
 XX PR 29-JUL-1997; 97US-0902615.  
 XX PR 01-APR-1997; 97US-0833457.  
 XX PR 24-JUN-1997; 97US-0881227.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX PA (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.  
 XX PI Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;  
 XX

```

DR WPI; 1998-542293/46.
DR N-PSDB; AAX14142.
XX
PT New isolated Helicobacter polynucleotides - used to develop products
PT for the diagnosis, prevention and treatment of Helicobacter
PT infections and gastrointestinal diseases
XX
PS Claim 8; Page 711-713; 2054pp; English.
XX
CC This sequence represents a Helicobacter pylori GHPO protein of the
CC invention. The polypeptides can be used for preventing or treating
CC Helicobacter infections, and gastroduodenal diseases associated with
CC these infections, including acute, chronic, and atrophic gastritis, and
CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be
CC used for the production of antibodies. The products can also be used for
CC detection and diagnosis.
XX
SQ Sequence 505 AA;

Query Match 100.0%; Score 157; DB 19; Length 505;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVDKVKQTAFGAPVWDNNVITAGPRG 29
Db 1 mvdkvkqtafgavpwddnnvitaqprg 29

RESULT 7
AAW54146
ID AAW54146 standard; Protein; 505 AA.
XX
AC AAW54146;
XX
DT 20-JUL-1998 (first entry)
XX
DE Helicobacter pylori isolate 921023 catalase.
XX
KW Treatment; prevention; gastroduodenal disease; detection.
XX
OS Helicobacter pylori.
XX
PN WO9806853-A1.
XX
PD 19-FEB-1998.
XX
XX
XX 14-AUG-1997; 97WO-AU00515.
XX
XX 15-AUG-1996; 96US-0695987.
XX
XX (CSLC-) CSL LTD.
XX
XX (UYNE-) UNIV NEW SOUTH WALES.
XX
XX Doldige CV, Hazell SL, Lee A, Radcliff FJ;
XX
XX WPI; 1998-159544/14.
XX
XX N-PSDB; AAV20911.
XX
XX
XX New isolated Helicobacter catalase nucleic acid - used to develop
XX products for the treatment or prevention of Helicobacter infection,
XX particularly H. pylori gastroduodenal disease
XX
XX Disclosure; Page 37-38; 46pp; English.
XX
XX
XX The Helicobacter pylori catalase is a protein of 58650 Da. The protein,
XX gene sequence and products, such as an immunogenic fragment, can be used
XX for the treatment or prevention of Helicobacter infection, particularly
XX H. pylori infections which cause gastroduodenal disease. They can also
XX be used for the detection and diagnosis of Helicobacter infection.
XX
XX Sequence 505 AA;
SQ

```

XX Protein identification; signal transduction pathway; metabolic pathway;  
KW hydriisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
OS  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123348.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 06-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0147935.  
PR 11-AUG-1999; 99US-0148171.  
PR 12-AUG-1999; 99US-0148319.  
PR 13-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.

PR	31-AUG-1999;	99US-0151438.	PN	EP1033405-A2.
PR	01-SEP-1999;	99US-0151930.	XX	
PR	07-SEP-1999;	99US-0152363.	PD	
PR	10-SEP-1999;	99US-0153070.	XX	06-SEP-2000.
PR	13-SEP-1999;	99US-0153758.	PF	25-FEB-2000; 2000EP-0301439.
PR	15-SEP-1999;	99US-0154018.	XX	
PR	16-SEP-1999;	99US-0154039.	PR	25-FEB-1999;
PR	20-SEP-1999;	99US-0154779.	PR	99US-0121825.
PR	22-SEP-1999;	99US-0155139.	PR	99US-0123180.
PR	23-SEP-1999;	99US-0155486.	PR	99US-0123548.
PR	24-SEP-1999;	99US-0155659.	PR	99US-0125788.
PR	28-SEP-1999;	99US-0156458.	PR	99US-0126264.
PR	29-SEP-1999;	99US-0156596.	PR	99US-0126765.
PR	04-OCT-1999;	99US-0157117.	PR	99US-0127462.
PR	05-OCT-1999;	99US-0157753.	PR	99US-0128234.
PR	06-OCT-1999;	99US-0157865.	PR	99US-0128714.
PR	07-OCT-1999;	99US-0158029.	PR	99US-0129845.
PR	08-OCT-1999;	99US-0158232.	PR	99US-0130077.
PR	12-OCT-1999;	99US-0158369.	PR	99US-0130449.
PR	13-OCT-1999;	99US-0159293.	PR	99US-0130510.
PR	13-OCT-1999;	99US-0159294.	PR	99US-0130891.
PR	13-OCT-1999;	99US-0159295.	PR	99US-0131449.
PR	14-OCT-1999;	99US-0159329.	PR	99US-0132048.
PR	14-OCT-1999;	99US-0159330.	PR	99US-0132407.
PR	14-OCT-1999;	99US-0159331.	PR	99US-0132484.
PR	14-OCT-1999;	99US-0159637.	PR	99US-0132485.
PR	14-OCT-1999;	99US-0159638.	PR	99US-0132486.
PR	18-OCT-1999;	99US-0159584.	PR	99US-0132487.
PR	21-OCT-1999;	99US-0160741.	PR	99US-0132863.
PR	21-OCT-1999;	99US-0160767.	PR	99US-0134256.
PR	21-OCT-1999;	99US-0160768.	PR	99US-0134218.
PR	21-OCT-1999;	99US-0160770.	PR	99US-0134219.
PR	21-OCT-1999;	99US-0160814.	PR	99US-0134221.
PR	21-OCT-1999;	99US-0160815.	PR	99US-0134370.
PR	22-OCT-1999;	99US-0160980.	PR	99US-0134768.
PR	22-OCT-1999;	99US-0160981.	PR	99US-0134941.
PR	22-OCT-1999;	99US-0160989.	PR	99US-0135124.
PR	25-OCT-1999;	99US-0161404.	PR	99US-0135353.
PR	25-OCT-1999;	99US-0161405.	PR	99US-0135629.
PR	25-OCT-1999;	99US-0161406.	PR	99US-0136021.
PR	26-OCT-1999;	99US-0161359.	PR	99US-0136392.
PR	26-OCT-1999;	99US-0161360.	PR	99US-0136782.
PR	26-OCT-1999;	99US-0161361.	PR	99US-0137222.
PR	28-OCT-1999;	99US-0161920.	PR	99US-0137528.
PR	28-OCT-1999;	99US-0161992.	PR	99US-0137502.
PR	28-OCT-1999;	99US-0161993.	PR	99US-0137724.
PR	29-OCT-1999;	99US-0162142.	PR	99US-0138094.
			PR	99US-0138540.
			PR	99US-0138847.
			PR	99US-0139119.
			PR	99US-0139452.
			PR	99US-0139453.
			PR	99US-0139454.
			PR	99US-0139492.
			PR	99US-0139454.
			PR	99US-0139455.
			PR	99US-0139456.
			PR	99US-0139457.
			PR	99US-0139458.
			PR	99US-0139459.
			PR	99US-0139460.
			PR	99US-0139461.
			PR	99US-0139462.
			PR	99US-0139463.
			PR	99US-0139750.
			PR	99US-0139763.
			PR	99US-0139817.
			PR	99US-0139899.
			PR	99US-0140353.
			PR	99US-0140354.
			PR	99US-0140695.
			PR	99US-0140823.
			PR	99US-0140891.
			PR	99US-0140991.
			PR	99US-0141287.
			PR	99US-0141842.

Query Match 47.8%; Score 75; DB 21; Length 377;  
 Best Local Similarity 61.9%; Pred. No. 0.004;  
 Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFCAPVWDNNVITAGPRG 29  
 || |||||:::|:| |||||  
 Db 18 ttmsgapvwnnssmtvgprg 38

RESULT 10  
 AAG23826  
 ID AAG23826 standard; Protein; 394 AA.  
 XX  
 AC AAG23826;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 27276.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX

PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147460.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 16-AUG-1999; 99US-0148684.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.

PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 22-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 47.8%; Score 75; DB 21; Length 394;  
Best Local Similarity 61.9%; Pred. No. 0.0042;  
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDNNVITAGPRG 29  
|| |||||::: :| ||||  
Db 35 ttmsgapvwnnssmtvgprg 55

RESULT 11

AAV18131  
ID AAY18131 standard; Protein; 492 AA.  
XX AAY18131;  
XX  
DT 11-AUG-1999 (first entry)  
XX  
DE Clone of A. thaliana strong light adapting protein.  
XX Strong light adapting condition; light-resistant plant; transgenic plant.  
XX Arabidopsis thaliana.  
XX JP11137253-A.  
PN  
XX  
PD 25-MAY-1999.  
XX  
PF 07-NOV-1997; 97JP-0306044.  
XX  
PR 07-NOV-1997; 97JP-0306044.  
XX  
PA (SUMO ) SUMITOMO CHEM CO LTD.

XX WPI; 1999-364702/31.  
DR N-PSDB; AAX77100.  
XX  
XX  
PT Induction of gene to be expressed in plant under a strong light  
PT adapting condition - useful for enabling plant to grow in desert  
XX  
XX PS Claim 7; Page 12-13; 16pp; Japanese.  
XX  
XX This sequence represents a clone of the protein of the invention.  
CC The invention relates to a gene that is induced and expressed in a plant  
CC under a strong light adapting conditions, and is prepared by a procedure  
CC in which one of the following light irradiations (1) to (3) is made on a  
CC plant already irradiated by a light of a light intensity of 60-100 mu  
CC E/sq. m s consisting of a continuous wave length component of a wave  
CC length region of 300-800 nm at 15 to 38 degrees C and the gene expression  
CC of which is induced in the cell of said plant: (1) A light of a light  
CC intensity of 300-700 mu E/sq. m s consisting of a continuous wave length  
CC component of a wave length region of 300-800 nm at 15-38 degrees C for  
CC 1 hour to several days; (2) a light of a light intensity of 200-600 mu  
CC E/sq. m s consisting of a continuous wave length component of a wave  
CC length region of 350-560 nm at 15-38 degrees C for several hours to  
CC several days; (3) a light of intensity 75-270 mu E/sq. m s consisting of  
CC a continuous wave length component of a wave length region of 300-800 nm  
CC at 0-10 degrees C for several hours to several days. The method can be  
CC used for generating a transgenic plant that is able to grow in the  
CC desert.  
XX Sequence 492 AA;  
SQ

Query Match 47.8%; Score 75; DB 20; Length 492;  
Best Local Similarity 61.9%; Pred. NO. 0.0054;  
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDNNVITAPRG 29  
|| |||||:::| ||||  
Db 18 ttmsgapvwnnsmtdvprg 38

RESULT 12  
AAG43132  
ID AAG43132 standard; Protein; 492 AA.  
XX  
XX AC AAG43132;  
XX  
XX DT 18-OCT-2000 (first entry)  
XX  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 53875.  
XX  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX OS Arabidopsis thaliana.  
XX  
XX PN EP1033405-A2.  
XX  
XX PD 06-SEP-2000.  
XX  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 23-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 03-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 26-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 47.8%; Score 75; DB 21; Length 492;  
Best Local Similarity 61.9%; Pred. No. 0.0054;  
Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAGCAPVWDNNVITAGPRG 29  
|| |||||:::| |||||  
Db 18 ttmsgapvwnnsmvtgprg 38

RESULT 13  
AAG43903  
ID AAG43903 standard; Protein; 512 AA.  
XX  
AC AAG43903;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 54930.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.





PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 47.8%; Score 75; DB 21; Length 512;  
 Best Local Similarity 61.9%; Pred. No. 0.0056;  
 Matches 13; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 9 TTAFGAPVWDDNNVITAGPRG 29  
 || |||||::: || ||||  
 Db 18 ttngagpvwnnnssmtvgprg 38

RESULT 14  
 AAY78358  
 ID AAY78358 standard; Protein; 483 AA.

AC AAY78358;  
 XX  
 XX 05-MAY-2000 (first entry)  
 XX  
 XX Bacillus thermoglucosidasius catalase kat 19 SEQ ID NO:9.  
 DE  
 XX Catalase; genetic engineering; hydrogen peroxide decomposition;  
 KW contact lens; disinfectant.  
 KW  
 XX Bacillus thermoglucosidasius.  
 OS  
 XX US6022721-A.  
 PN  
 XX 08-FEB-2000.  
 PD  
 XX 20-FEB-1998; 98US-0027166.  
 PF  
 XX 03-JAN-1997; 97TW-0100018.  
 PR  
 XX 11-DEC-1997; 97CN-0120386.  
 XX  
 PA (BIOT-) DEV CENT BIOTECHNOLOGY.

XX Hwong C, Lo C, Chang LE;  
 XX WPI; 2000-160378/14.  
 XX N-PSDB; AA288398.  
 XX  
 PT Isolated catalase gene derived from Bacillus thermoglucosidasius useful  
 PT for decomposing hydrogen peroxide in residual disinfectant remaining on  
 PT a contact lens.  
 XX  
 PS Disclosure; Column 25-28; 30pp; English.  
 XX  
 CC The present sequence represents a catalase isolated from Bacillus  
 CC thermoglucosidasius. A composition containing a B. thermoglucosidasius  
 CC catalase is used to decompose hydrogen peroxide present in residual  
 CC disinfectant remaining on a contact lens to avoid undesired injury to  
 CC users. The catalase enzyme can also be used to treat textile substances  
 CC which have been bleached with hydrogen peroxide. The catalase is derived  
 CC from a non-mammalian source and reducing the need to use bovine liver  
 CC catalases so the risk of cross infection from cows to humans is reduced.  
 CC The production procedure is simple and has low cost with high yield of  
 CC enzyme.

XX SQ Sequence 483 AA;  
 Query Match 45.2%; Score 71; DB 21; Length 483;  
 Best Local Similarity 60.9%; Pred. No. 0.021;  
 Matches 14; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
 QY 7 KOTAFGAPVWDDNNVITAGPRG 29  
 | |||::||| | | : ||| ||  
 Db 5 kltswgapvghqnsmtagsr 27  
 RESULT 15  
 AAG19059  
 ID AAG19059 standard; Protein; 115 AA.  
 XX  
 AC AAG19059;  
 XX  
 XX 17-OCT-2000 (first entry)  
 DT  
 XX  
 XX Zea mays protein fragment SEQ ID NO: 20715.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 OS Zea mays subsp. mays.  
 XX  
 XX EP1033405-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 09-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 11-JUN-1999; 99US-0139119.  
PR 12-JUN-1999; 99US-0139452.  
PR 13-JUN-1999; 99US-0139453.  
PR 14-JUN-1999; 99US-0139453.  
PR 15-JUN-1999; 99US-0139453.  
PR 16-JUN-1999; 99US-0139454.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 19-JUN-1999; 99US-0139456.  
PR 20-JUN-1999; 99US-0139457.  
PR 21-JUN-1999; 99US-0139458.  
PR 22-JUN-1999; 99US-0139459.  
PR 23-JUN-1999; 99US-0139460.  
PR 24-JUN-1999; 99US-0139461.  
PR 25-JUN-1999; 99US-0139462.  
PR 26-JUN-1999; 99US-0139463.  
PR 27-JUN-1999; 99US-0139750.  
PR 28-JUN-1999; 99US-0139763.  
PR 29-JUN-1999; 99US-0139817.  
PR 30-JUN-1999; 99US-0139899.  
PR 31-JUN-1999; 99US-0140353.  
PR 01-JUL-1999; 99US-0140354.  
PR 02-JUL-1999; 99US-0140695.  
PR 03-JUL-1999; 99US-0140823.  
PR 04-JUL-1999; 99US-0140991.  
PR 05-JUL-1999; 99US-0141287.  
PR 06-JUL-1999; 99US-0141842.  
PR 07-JUL-1999; 99US-0142154.  
PR 08-JUL-1999; 99US-0142055.  
PR 09-JUL-1999; 99US-0142390.  
PR 10-JUL-1999; 99US-0142803.  
PR 11-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 17-JUL-1999; 99US-0144086.  
PR 18-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 20-JUL-1999; 99US-0144332.  
PR 21-JUL-1999; 99US-0144333.  
PR 22-JUL-1999; 99US-0144334.  
PR 23-JUL-1999; 99US-0144335.  
PR 24-JUL-1999; 99US-0144352.  
PR 25-JUL-1999; 99US-0144632.  
PR 26-JUL-1999; 99US-0144884.  
PR 27-JUL-1999; 99US-0144814.  
PR 28-JUL-1999; 99US-0145086.  
PR 29-JUL-1999; 99US-0145088.  
PR 30-JUL-1999; 99US-0145085.  
PR 01-AUG-1999; 99US-0145087.  
PR 02-AUG-1999; 99US-0145089.  
PR 03-AUG-1999; 99US-0145192.  
PR 04-AUG-1999; 99US-0145145.  
PR 05-AUG-1999; 99US-0145218.  
PR 06-AUG-1999; 99US-0145224.  
PR 07-AUG-1999; 99US-0145276.  
PR 08-AUG-1999; 99US-0145913.  
PR 09-AUG-1999; 99US-0145918.  
PR 10-AUG-1999; 99US-0145919.  
PR 11-AUG-1999; 99US-0145951.  
PR 12-AUG-1999; 99US-0146386.  
PR 13-AUG-1999; 99US-0146388.  
PR 14-AUG-1999; 99US-0147038.  
PR 15-AUG-1999; 99US-0147038.  
PR 16-AUG-1999; 99US-0147204.  
PR 17-AUG-1999; 99US-0147302.  
PR 18-AUG-1999; 99US-0147192.  
PR 19-AUG-1999; 99US-0147260.  
PR 20-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.  
PR 07-AUG-1999; 99US-0147493.  
PR 08-AUG-1999; 99US-0147935.  
PR 09-AUG-1999; 99US-0148171.  
PR 10-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148341.  
PR 12-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 14-AUG-1999; 99US-0149368.  
PR 15-AUG-1999; 99US-0149175.  
PR 16-AUG-1999; 99US-0149426.  
PR 17-AUG-1999; 99US-0149722.  
PR 18-AUG-1999; 99US-0149723.  
PR 19-AUG-1999; 99US-0149929.  
PR 20-AUG-1999; 99US-0149929.  
PR 21-AUG-1999; 99US-0149929.  
PR 22-AUG-1999; 99US-0149930.  
PR 23-AUG-1999; 99US-0150566.  
PR 24-AUG-1999; 99US-0150884.  
PR 25-AUG-1999; 99US-0151065.  
PR 26-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 28-AUG-1999; 99US-0151303.  
PR 29-AUG-1999; 99US-0151438.  
PR 30-AUG-1999; 99US-0151930.  
PR 01-SEP-1999; 99US-0152363.  
PR 02-SEP-1999; 99US-0153070.  
PR 03-SEP-1999; 99US-0153758.  
PR 04-SEP-1999; 99US-0154018.  
PR 05-SEP-1999; 99US-0154039.  
PR 06-SEP-1999; 99US-0154779.  
PR 07-SEP-1999; 99US-0155139.  
PR 08-SEP-1999; 99US-0155486.  
PR 09-SEP-1999; 99US-0155659.  
PR 10-SEP-1999; 99US-0156458.  
PR 11-SEP-1999; 99US-0156596.  
PR 12-SEP-1999; 99US-0157117.  
PR 13-SEP-1999; 99US-0157753.  
PR 14-SEP-1999; 99US-0157865.  
PR 15-SEP-1999; 99US-0158029.  
PR 16-SEP-1999; 99US-0158232.  
PR 17-SEP-1999; 99US-0158369.  
PR 18-SEP-1999; 99US-0159293.  
PR 19-SEP-1999; 99US-0159294.  
PR 20-SEP-1999; 99US-0159295.  
PR 21-SEP-1999; 99US-0159329.  
PR 22-SEP-1999; 99US-0159330.  
PR 23-SEP-1999; 99US-0159331.  
PR 24-SEP-1999; 99US-0159637.  
PR 25-SEP-1999; 99US-0159638.  
PR 26-SEP-1999; 99US-0159584.  
PR 27-SEP-1999; 99US-0160741.  
PR 28-SEP-1999; 99US-0160767.  
PR 29-SEP-1999; 99US-0160768.  
PR 30-SEP-1999; 99US-0160770.  
PR 01-OCT-1999; 99US-0160814.  
PR 02-OCT-1999; 99US-0160815.  
PR 03-OCT-1999; 99US-0160815.  
PR 04-OCT-1999; 99US-0160980.  
PR 05-OCT-1999; 99US-0160981.  
PR 06-OCT-1999; 99US-0160981.  
PR 07-OCT-1999; 99US-0160981.  
PR 08-OCT-1999; 99US-0160981.  
PR 09-OCT-1999; 99US-0161404.  
PR 10-OCT-1999; 99US-0161405.  
PR 11-OCT-1999; 99US-0161406.  
PR 12-OCT-1999; 99US-0161359.  
PR 13-OCT-1999; 99US-0161360.  
PR 14-OCT-1999; 99US-0161361.  
PR 15-OCT-1999; 99US-0161361.  
PR 16-OCT-1999; 99US-0161920.  
PR 17-OCT-1999; 99US-0161992.  
PR 18-OCT-1999; 99US-0161993.  
PR 19-OCT-1999; 99US-0162142.

Query Match 44.6%; Score 70; DB 21; Length 115;  
Best Local Similarity 57.1%; Pred. No. 0.006;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 9 TTAAGGAPVWDDNNVITAGPRG 29  
Db 18 ttmsgapvwnnsaltvggrg 38

RESULT 16  
AAG28201  
ID AAG28201 standard; Protein: 492 AA.  
AC AAG28201;  
XX  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33331.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
OS Arabidopsis thaliana.  
PN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126765.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 05-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.



PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 08-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 43.3%; Score 68; DB 21; Length 536;  
Best Local Similarity 57.1%; Pred. No. 0.069;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFGAPVDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 62 ttnsgapvwnnssltvgtgr 82

RESULT 18  
AAW40233  
ID AAW40233 standard; peptide; 15 AA.  
XX  
AC AAW40233;  
XX  
DT 08-JUN-1998 (first entry)  
XX

```

Query Match          42.7%; Score 67; DB 19; Length 15;
Best Local Similarity 92.9%; Pred. No. 0.0017;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 VNKDVKQTAFGAP 15
        ||||| ||||
Db       2 vnkdkvqtxafg 15

RESULT 19
ABB30447
ID      ABB30447 standard; Peptide; 57 AA.
XX
XX      ABB30447;
XX
XX      01-FEB-2002 (first entry)
DT
XX
XX      Peptide #3098 encoded by breast cell single exon nucleic acid probe.
DE
XX      Human; microarray; single exon probe; gene expression; breast;
KW      disease; cancer.
KW
XX
XX      Homo sapiens.
OS
XX
XX      WO200157271-A2.
PN
XX
XX      09-AUG-2001.
PD
XX
XX      30-JAN-2001; 2001WO-US00662.
PF
XX
XX      04-FEB-2000; 2000US-0180312.
PR
PR      26-MAY-2000; 2000US-0207456.

```

Query Match	42.0%;	Score 66;	DB 22;	Length 57;
Best Local Similarity	66.7%;	Pred. No. 0.011;		
Matches	14;	Conservative	0;	Mismatches 7; Indels 0; Gaps 0;
Qy	9	TTAFGAPVWDDNNVITAGPRG	29	
Db	6	ttgagnpvgdklnvitvprg	26	
RESULT	20			
ABB35615				
ID	ABB35615	standard;	Peptide; 57	AA.
XX	XX			
XX	ABB35615;			
XX	XX			
DT	04-FEB-2002	(first entry)		
XX				
DE	Peptide #3121	encoded by human foetal liver single exon probe.		
XX				
KW	Human;	foetal liver;	gene expression;	single exon nucleic acid probe.
XX				
OS	Homo sapiens.			
XX				
PN	WO200157277-A2.			
XX				
PD	09-AUG-2001.			
XX				
PF	30-JAN-2001;	2001WO-US00669.		
XX				
PR	04-FEB-2000;	2000US-0180312.		
PR	26-MAY-2000;	2000US-0207456.		
PR	30-JUN-2000;	2000US-0608408.		

PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
PT  
XX  
XX Claim 27; SEQ ID NO 28250; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 57 AA;  
SQ

Query Match 42.0%; Score 66; DB 22; Length 57;  
Best Local Similarity 66.7%; Pred. No. 0.011; 7; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAAGAPVDDNNVITAGPRG 29  
II III III III  
Db 6 ttgagnpvgdklnvityprg 26

RESULT 21  
ABB21044  
ID ABB21044 standard; Protein; 57 AA.  
XX  
XX ABB21044;  
XX  
XX 23-JAN-2002 (first entry)  
DT  
XX  
XX Protein #3043 encoded by probe for measuring heart cell gene expression.  
XX  
XX Human; gene expression; heart; microarray; vascular system;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157274-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US00666.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR  
XX  
XX 26-MAY-2000; 2000US-0207456.  
PR  
XX  
XX 30-JUN-2000; 2000US-0608408.  
PR  
XX  
XX 03-AUG-2000; 2000US-0632366.  
PR  
XX  
XX 21-SEP-2000; 2000US-0234687.  
PR  
XX  
XX 27-SEP-2000; 2000US-0236359.  
PR  
XX  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX  
XX WPI; 2001-488899/53.  
DR

XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
XX  
XX Claim 15; SEQ ID NO 22814; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting, the  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 57 AA;  
SQ

Query Match 42.0%; Score 66; DB 22; Length 57;  
Best Local Similarity 66.7%; Pred. No. 0.011; 7; Indels 0; Gaps 0;  
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAAGAPVDDNNVITAGPRG 29  
II III III III  
Db 6 ttgagnpvgdklnvityprg 26

RESULT 22  
AAM56432  
ID AAM56432 standard; Protein; 57 AA.  
XX  
XX AAM56432;  
XX  
XX 05-NOV-2001 (first entry)  
DT  
XX  
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 28537.  
XX  
XX Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200157275-A2.  
PN  
XX  
XX 09-AUG-2001.  
PD  
XX  
XX 30-JAN-2001; 2001WO-US00667.  
PF  
XX  
XX 04-FEB-2000; 2000US-0180312.  
PR  
XX  
XX 26-MAY-2000; 2000US-0207456.  
PR  
XX  
XX 30-JUN-2000; 2000US-0608408.  
PR  
XX  
XX 03-AUG-2000; 2000US-0632366.  
PR  
XX  
XX 21-SEP-2000; 2000US-0234687.  
PR  
XX  
XX 27-SEP-2000; 2000US-0236359.  
PR  
XX  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI  
XX  
XX WPI; 2001-483446/52.  
DR  
XX  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
XX  
XX Example 4; SEQ ID NO: 28537; 650pp + Sequence Listing; English.  
PS  
XX  
XX The present invention provides a number of single exon nucleic acid  
CC



CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 57 AA;

Query Match 42.0%; Score 66; DB 22; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 0.011;  
 Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTAAGAPVWDDNNVITAGPRG 29  
 || | || | || | || | || |  
 Db 6 ttgagnpvgdklnvitvgrg 26

RESULT 23  
 AAM68813  
 ID AAM68813 standard; Protein; 57 AA.  
 XX  
 AC AAM68813;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 29119.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157276-A2.  
 PN  
 XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for  
 analyzing gene expression in human bone marrow -

Example 4; SEQ ID NO: 29119; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX

SQ Sequence 57 AA;

Query Match 42.0%; Score 66; DB 22; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 0.011;  
 Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTAAGAPVWDDNNVITAGPRG 29  
 || | || | || | || | || |  
 Db 6 ttgagnpvgdklnvitvgrg 26

RESULT 24  
 AAM16631  
 ID AAM16631 standard; Protein; 57 AA.  
 XX  
 AC AAM16631;  
 XX

DT 12-OCT-2001 (first entry)

XX Peptide #3065 encoded by probe for measuring cervical gene expression.  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157278-A2.  
 PN  
 XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488901/53.

Human genome-derived single exon nucleic acid probes useful for  
 analyzing gene expression in human cervical epithelial cells -

Claim 27; SEQ ID NO 21457; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes  
 CC (SENPs: see AAI10068-AAI28459). The present sequence is a peptide encoded  
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.

CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 57 AA;

Query Match 42.0%; Score 66; DB 22; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 0.011;  
 Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTAAGAPVWDDNNVITAGPRG 29  
 || | || | || | || | || |  
 Db 6 ttgagnpvgdklnvitvgrg 26

RESULT 25  
 AAM29117  
 ID AAM29117 standard; Protein; 57 AA.  
 XX

AC AAM29117;  
 XX 17-OCT-2001 (first entry)  
 XX Peptide #3154 encoded by probe for measuring placental gene expression.  
 DE Probe; microarray; human; placenta; antenatal diagnosis;  
 XX genetic disorder.  
 KW Homo sapiens.  
 XX WO200157272-A2.  
 PN 09-AUG-2001.  
 XX 30-JAN-2001; 2001WO-US00663.  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488897/53.  
 XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX Claim 27; SEQ ID No 29386; 654pp; English.  
 XX The present invention relates to single exon nucleic acid probes (SENP;  
 CC see AA131315-AA157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 XX Sequence 57 AA;  
 SQ Query Match 42.0%; Score 66; DB 22; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 0.011; Mismatches 0; Gaps 0;  
 Matches 14; Conservative 0; Indels 7; Indels 0; Gaps 0;  
 QY 9 TTAFGAPVWDNNVITAGPRG 29  
 Db || || | ||| ||||  
 6 ttgagnpvgdklnvltgprg 26  
 RESULT 26  
 AAM04348  
 ID AAM04348 standard; Protein; 57 AA.  
 AC AAM04348;  
 XX 09-OCT-2001 (first entry)  
 DT Peptide #3030 encoded by probe for measuring breast gene expression.  
 XX Probe; human; breast disease; breast cancer; development disorder;  
 DE inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 KW Homo sapiens.  
 XX WO200157270-A2.  
 PN 09-AUG-2001.  
 XX 29-JAN-2001; 2001WO-US00661.  
 XX 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-476286/51.  
 XX Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX Claim 27; SEQ ID No 13088; 322pp; English.  
 XX The present invention relates to novel single exon nucleic acid probes  
 CC (see AA100010-AA110067). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for measuring human gene expression in  
 CC a human breast sample, where the probe hybridizes at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX Sequence 57 AA;  
 SQ Query Match 42.0%; Score 66; DB 22; Length 57;  
 Best Local Similarity 66.7%; Pred. No. 0.011; Mismatches 0; Gaps 0;  
 Matches 14; Conservative 0; Indels 7; Indels 0; Gaps 0;  
 QY 9 TTAFGAPVWDNNVITAGPRG 29  
 Db || || | ||| ||||  
 6 ttgagnpvgdklnvltgprg 26  
 RESULT 27  
 AAY44174  
 ID AAY44174 standard; Protein; 483 AA.  
 XX AAY44174;  
 AC AAY44174;  
 XX 01-FEB-2000 (first entry)  
 DT Bacillus thermoglucosidasius catalase protein.  
 XX Catalase; PCR; primer; amplification; expression; E.coli;  
 DE decomposition; hydrogen peroxide; disinfectant; contact lens.  
 KW Bacillus thermoglucosidasius.  
 XX JP11243961-A.  
 PN 14-SEP-1999.  
 XX 06-MAR-1998; 98JP-0055299.  
 XX 06-MAR-1998; 98JP-0055299.  
 PR (BIOT-) DEV CENT BIOTECHNOLOGY.

```

CC enzyme.
XX SQ Sequence 483 AA;
Query Match 42.0%; Score 66; DB 21; Length 483;
Best Local Similarity 60.9%; Pred. No. 0.12;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 7 KQTAFGAPVWDDNNVITAGPRG 29
Dd 5 klttswgapvgnqnsitagpqp 27
      | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | |

RESULT 29
AAB46776
AAB46776 standard; Protein; 509 AA.
XX AC AAB46776;
XX DT 19-APR-2001 (first entry)
XX DE V. rumoiensis S-1 catalase protein SEQ ID 2.
XX KW Catalase; hydrogen peroxide decomposition; waste water.
XX OS Vibrio rumoiensis.
XX PN JP2000316584-A.
XX PD 21-NOV-2000.
XX PF 14-MAY-1999; 99JP-0134831.
XX PR 14-MAY-1999; 99JP-0134831.
XX PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX DR WPI; 2001-106416/12.
XX DR N-PSDB; AAF25916.
XX PT New Vibrio rumoiensis S-1 catalase polypeptide and polynucleotide for
XX PT treating waste water -
XX PS Claim 2; Page 7-8; 10pp; Japanese.
XX CC This invention describes a novel Vibrio rumoiensis S-1 catalase
XX CC polynucleotide which comprises the sequence (1) given in the
XX CC specification. The invention also describes (1) a catalase polypeptide
XX CC comprising at least residues 127 to 509 of the fully defined amino acid
XX CC given in the specification, (2) a recombinant vector carrying comprising
XX CC the gene; and (3) a transformant transformed by the above recombinant
XX CC vector. The catalase is useful as an enzymatic agent decomposing hydrogen
XX CC peroxide contained in waste water.
XX SQ Sequence 509 AA;

Query Match 42.0%; Score 66; DB 22; Length 509;
Best Local Similarity 61.9%; Pred. No. 0.13;
Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFGAPVWDDNNVITAGPRG 29
Dd 14 ttdfgapvvtndrsitagprg 34
      | | | | | | | | | | | | | | | |
      | | | | | | | | | | | | | | | |

RESULT 30
AAP82004
ID AAP82004 standard; protein; 527 AA.
XX AC AAP82004;
XX XX

```

DT 15-OCT-1990 (first entry)  
 XX Recombinant catalase.  
 DE Catalase; inflammation; antiinflammatory drug.  
 KW  
 XX  
 KW Homo sapiens.  
 XX  
 XX JP63017693-A.  
 XX  
 XX 25-JAN-1988.  
 XX  
 XX 06-JUL-1986; 86JP-0159690.  
 XX  
 XX 06-JUL-1986; 86JP-0159690.  
 XX  
 XX (MITK ) MITSUI TOATSU CHEM IND.  
 XX  
 XX WPI; 1988-061234/09.  
 DR N-PSDB; AAN82012.  
 XX  
 XX Human catalase gene - used to produce catalase  
 PT for antiinflammatory drugs.  
 PT  
 XX  
 PS Claim 1; Page 1; 5pp; Japanese.  
 XX  
 XX The DNA encoding the gene can be used to produce recombinant  
 CC catalase which, with superoxide dismutase, is an important anti  
 CC inflammatory factor in vivo and can be used as a drug.  
 CC  
 CC See also AAN82011.  
 XX  
 SQ Sequence 527 AA;

Query Match 42.0%; Score 66; DB 9; Length 527;  
 Best Local Similarity 66.7%; Pred. No. 0.14;  
 Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
 || ||| |||| ||||  
 Db 28 ttgagnpvgdklnvitygprg 48

RESULT 31  
 AAW54361  
 ID AAW54361 standard; protein; 527 AA.

XX AC AAW54361;  
 XX  
 XX 14-AUG-1998 (first entry)  
 XX  
 DE Catalase.  
 XX  
 XX Endometrium; hyperplasia; adenocarcinoma; proliferative phase;  
 KW 2D gel electrophoresis; detection.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO9810291-A1.  
 XX  
 XX 12-MAR-1998.  
 XX  
 XX 05-SEP-1997; 97WO-GB02394.  
 XX  
 XX 08-APR-1997; 97GB-0007132.  
 PR  
 XX 06-SEP-1996; 96GB-0018600.  
 XX  
 XX (CLIN-) CENT CLINICAL & BASIC RES.  
 PA  
 XX Byrjalsen I, Fey SJ, Larsen P;  
 PI  
 XX WPI; 1998-207057/18.  
 DR  
 XX

PT Biochemical markers of human endometrium - useful for, e.g.  
 PT diagnosis of hyperplasia and adenocarcinoma  
 XX  
 PS Disclosure; Page 23; 77pp; English.  
 XX  
 CC Proteins AAW54349-W54364 are examples of proteins produced in the  
 CC endometrium during the hyperplasia, adenocarcinoma or proliferative  
 CC phase of the endometrium. The presence and quantities of these proteins  
 CC can be detected using 2D gel electrophoresis comparison of cell lysates.  
 CC The proteins can be used as biochemical markers to detect the phase of  
 CC the endometrium and can be measured in body fluids, obviating the need  
 CC for endometrial biopsies.  
 XX  
 SQ Sequence 527 AA;

Query Match 42.0%; Score 66; DB 19; Length 527;  
 Best Local Similarity 66.7%; Pred. No. 0.14;  
 Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
 || ||| |||| ||||  
 Db 28 ttgagnpvgdklnvitygprg 48

RESULT 32  
 AAW96321  
 ID AAW96321 standard; protein; 527 AA.

XX AC AAW96321;  
 XX  
 XX 28-JUN-1999 (first entry)  
 XX  
 DE Human catalase sequence.  
 XX  
 XX Manganese containing superoxide dismutase; MnSOD; IDDM;  
 KW diabetes mellitus; treatment; therapy; nitric oxide; NO; beta cell;  
 KW fatty acid; lipotoxic; cytotoxic; cytokine; osteoporosis;  
 KW inflammatory disease; autoimmune disease; neurodegenerative disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO9906059-A2.  
 XX  
 XX 11-FEB-1999.  
 XX  
 XX 30-JUL-1998; 98WO-US15781.  
 XX  
 XX 03-MAR-1998; 98US-0055092.  
 PR  
 XX 30-JUL-1997; 97US-0055092.  
 XX  
 XX (BETA-) BETAGENE INC.  
 PA (TEXA ) UNIV TEXAS SYSTEM.  
 XX  
 XX Clark SA, Hohmeier H, Koyama K, Lee Y, Newgard CB;  
 PI Ohneda M, Shimabukuro, Thigpen A, Unger RH;  
 XX  
 XX WPI; 1999-153448/13.  
 DR  
 XX N-PSDB; AAX08433.  
 XX

PT Protection of mammalian cells against immunotoxicity or lipotoxicity  
 PT - used for treating, e.g. diabetes, obesity, wasting syndromes,  
 PT osteoporosis, inflammatory diseases, autoimmune diseases or  
 PT neurodegenerative diseases  
 XX  
 PS Disclosure; Page 242-244; 253pp; English.  
 XX  
 CC Inhibition of cytokine mediated immunotoxicity of cells can be  
 CC achieved by blocking free radical production or the accumulation of  
 CC free radicals in that cell. Treatment of insulin dependent diabetes  
 CC mellitus (IDDM) can be achieved by blocking nitric oxide (NO)  
 CC production in a pancreatic beta cell and by providing a composition  
 CC comprising an agent that reduces levels of fatty acids in the cells



CC the encoded polypeptides (AAI38642-AAI42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities, chemotactic/chemokinetic activity, haemostatic  
CC Activin/inhibin activity, cancer diagnosis and therapy, drug screening,  
CC and thrombolytic activity, cancer diagnosis and inflammation, leukaemias and  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: the sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 527 AA;

Query Match 42.0%; Score 66; DB 22; Length 527;  
Best Local Similarity 66.7%; Pred. No. 0.14;  
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29  
Db 28 ttgagpvgdkinvitvgrg 48  
|| | | | | | | | | | |

RESULT 35  
AAI41654  
ID AAI41654 standard; Protein; 558 AA.  
AC AAI41654;  
XX  
XX  
XX 22-OCT-2001 (first entry)  
XX  
DE Human polypeptide SEQ ID NO 6585.

XX Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.

XX Homo sapiens.  
XX  
XX WO200153312-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 26-DEC-2000; 2000WO-US34263.  
XX  
XX 21-JAN-2000; 2000US-0488725.  
XX 25-APR-2000; 2000US-0552317.  
XX 09-JUL-2000; 2000US-0598042.  
XX 19-JUL-2000; 2000US-0620312.  
XX 03-AUG-2000; 2000US-0653450.  
XX 14-SEP-2000; 2000US-0662191.  
XX 19-OCT-2000; 2000US-0693036.  
XX 29-NOV-2000; 2000US-0727344.

XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX  
XX WPI; 2001-442253/47.  
XX N-PSDB; AAI60810.  
XX

XX Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as; central nervous system injuries -

XX  
PS Example 2; SEQ ID NO 6585; 10078pp; English.  
XX  
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
CC the encoded polypeptides (AAI38642-AAI42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities, chemotactic/chemokinetic activity, haemostatic  
CC Activin/inhibin activity, cancer diagnosis and therapy, drug screening,  
CC and thrombolytic activity, cancer diagnosis and inflammation, leukaemias and  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: the sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 558 AA;

Query Match 42.0%; Score 66; DB 22; Length 558;  
Best Local Similarity 66.7%; Pred. No. 0.15;  
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29  
Db 59 ttgagpvgdkinvitvgrg 79  
|| | | | | | | | | | |

RESULT 36  
AAI79424  
ID AAI79424 standard; Protein; 565 AA.  
XX  
XX AAI79424;  
XX  
XX 06-NOV-2001 (first entry)  
XX Human protein SEQ ID NO 3070.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.

XX Homo sapiens.  
XX  
XX WO200157190-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 05-FEB-2001; 2001WO-US04098.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
XX 27-APR-2000; 2000US-0560875.  
XX 20-JUN-2000; 2000US-0598075.  
XX 19-JUL-2000; 2000US-0620325.  
XX 01-SEP-2000; 2000US-0654936.  
XX 15-SEP-2000; 2000US-0663561.  
XX 20-OCT-2000; 2000US-0693325.  
XX 30-NOV-2000; 2000US-0728422.

XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
XX  
XX WPI; 2001-476283/51.  
XX N-PSDB; AAI52557.

Query Match 38.9%; Score 61; DB 21; Length 499;

Best Local Similarity 44.08; Pred. No. 0.75;		Matches 11; Conservative 7; Mismatches 5; Indels 2; Gaps 1;
QY	6 VKQTTAFGAPVWDD--NNVITAGPR 28	
Db	434 lkqtsfgrpmwdnefsriatahpk 458	
RESULT 39		
AAAY71121		
ID	AAAY71121 standard; Protein; 563 AA.	
XX	AC	AAAY71121;
XX	XX	08-SEP-2000 (first entry)
XX	XX	Rat mitogenic regulator mox1.
XX	XX	Rat; mitogenic regulator; mox1; mitogenic oxidase; p65mox;
KW	KW	superoxide; reactive oxygen intermediate; ROI; cell division;
KW	KW	cytostatic; antipsoriatic; cardiant; antiarteriosclerotic; vasotropic;
KW	KW	antiangiogenic; hypotensive; drug development; treatment; cancer;
KW	KW	abnormal growth; psoriasis; prostatic hypertrophy; proliferation;
KW	KW	blood vessel; lymphatic vessel; benign prostatic hypertrophy;
KW	KW	cardiovascular disease; arteriovenous malformation; eye disorder;
KW	KW	hypertension; atherosclerosis; restenosis; angioplasty.
OS	OS	Rattus sp.
XX	XX	WO200028031-A2.
XX	XX	18-MAY-2000.
XX	XX	10-NOV-1999; 99WO-US26592.
XX	XX	10-NOV-1998; 98US-0107911.
PR	PR	17-AUG-1999; 99US-0149332.
PR	PR	27-AUG-1999; 99US-0151242.
XX	XX	(UYEM-) UNIV EMORY.
XX	XX	Lambeth JD, Lassegue BP, Griendling KK, Arnold RS, Guangjie C;
XX	XX	WPI; 2000-376545/32.
DR	DR	N-PSDB; AAD00694.
XX	XX	Protein capable of stimulating superoxide production, useful for
PT	PT	treating conditions associated with abnormal growth, including cancer
PT	PT	.
XX	XX	Claim 2; Page 98-100; 141pp; English.
PS	PS	The present sequence is a rat mitogenic oxidase mox1.
XX	XX	also referred as p65mox, which is capable of stimulating production of
CC	CC	superoxide, a reactive oxygen intermediate (ROI) that affects cell
CC	CC	division. The present sequence was obtained from a rat aortic
CC	CC	smooth muscle cell. The mox1 protein functions as a mitogenic regulator
CC	CC	and shows homology to human mox1 protein. It is expressed in aortic
CC	CC	smooth muscle cells and induced by angiotensin II, platelet-derived
CC	CC	growth factor (PDGF) and phorbol myristic acid (PMA).
CC	CC	The present sequence is useful in developing
CC	CC	drugs and therapies for treatment of conditions associated with abnormal
CC	CC	growth, including cancer, psoriasis, prostatic hypertrophy, benign
CC	CC	prostatic hypertrophy, cardiovascular disease, proliferation of vessels,
CC	CC	e.g. blood vessels and lymphatic vessels, arteriovenous malformation,
CC	CC	vascular problems associated with eye, atherosclerosis, hypertension,
CC	CC	and restenosis following angioplasty.
XX	XX	Sequence 563 AA;
Query Match 38.9%; Score 61; DB 21; Length 563;		
Best Local Similarity 44.08; Pred. No. 0.86;		

---

Best Local Similarity 44.08; Pred. No. 0.75;		Matches 11; Conservative 7; Mismatches 5; Indels 2; Gaps 1;
QY	6 VKQTTAFGAPVWDD--NNVITAGPR 28	
Db	498 lkqtsfgrpmwdnefsriatahpk 522	
RESULT 40		
AAAY73954		
ID	AAAY73954 standard; Protein; 152 AA.	
XX	AC	AAAY73954;
XX	XX	14-MAR-2000 (first entry)
XX	XX	Human prostate tumor EST fragment derived protein #141.
XX	XX	Pancreas; tumor; EST; expressed sequence tag; human; cytostatic;
KW	KW	treatment.
OS	OS	Homo sapiens.
XX	XX	DEL9820190-A1.
XX	XX	04-NOV-1999.
XX	XX	28-APR-1998; 98DE-1020190.
XX	XX	28-APR-1998; 98DE-1020190.
XX	XX	(META-) METAGEN GES GENOMFORSCHUNG MBH.
XX	XX	Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
DR	DR	WPI; 1999-621386/54.
DR	DR	N-PSDB; AAZ52904.
XX	XX	New human nucleic acid sequences from pancreatic tumors, and related
PT	PT	proteins
XX	XX	Claim 23; Page 368; 502pp; German.
PS	PS	This invention describes novel polypeptides and their encoding nucleic
XX	XX	acids derived from human pancreatic tumor tissue which have cytostatic
CC	CC	activity. The sequences are also useful in producing pharmaceutical
CC	CC	compositions for treatment of pancreatic tumors. AAY73814-Y74252
CC	CC	represent protein fragments encoded by the human pancreatic tumor cDNA
CC	CC	library derived expressed sequence tag (EST) sequences represented in
CC	CC	AAZ52858-Z53014.
XX	XX	Sequence 152 AA;
Query Match 36.9%; Score 58; DB 20; Length 152;		
Best Local Similarity 45.0%; Pred. No. 0.55;		
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;		
QY	8 QTTAFGAPVWDDNNVITAGP 27	
Db	52 rtqvfhapvmesgiltvvp 71	
Search completed: August 22, 2002, 07:44:07		
Job time: 66 sec		



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:44:32 ; Search time 21.52 Seconds  
(without alignments)  
32.916 Million cell updates/sec

Title: US-09-488-737-2  
Perfect score: 157  
Sequence: 1 MVNKKVQTTAFGAPVMDNNVITAGPRG 29

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Listing first 100 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	100.0	505	3	US-08-695-987-4
2	157	100.0	505	3	US-08-657-868B-4
3	157	100.0	505	3	US-08-657-868B-5
4	157	100.0	505	3	US-08-657-868B-6
5	153	97.5	505	3	US-08-695-987-2
6	73	46.5	15	3	US-08-695-987-6
7	71	45.2	483	3	US-09-027-166-9
8	66	42.0	483	3	US-09-027-166-7
9	66	42.0	527	4	US-09-126-109-10
10	57	36.3	495	4	US-08-980-241-5
11	56	35.7	527	2	US-08-365-486A-26
12	56	35.7	527	4	US-08-880-342-26
13	52.5	33.4	730	1	US-07-846-181-5
14	52.5	33.4	730	1	US-07-845-989-5
15	50	31.8	1188	1	US-08-201-697-4
16	49	31.2	676	4	US-09-061-768A-2
17	44.5	28.3	271	2	US-08-852-743-4
18	44.5	28.3	271	3	US-09-185-370-4
19	44	28.0	60	4	US-08-689-411-8
20	44	28.0	209	3	US-08-464-052-4
21	44	28.0	209	4	US-08-461-002-4
22	44	28.0	209	4	US-08-689-411-4
23	44	28.0	488	1	US-07-794-393-2
24	44	28.0	488	1	US-08-001-711-2
25	44	28.0	488	3	US-08-704-711A-22
26	44	28.0	489	4	US-08-448-489-11
27	44	28.0	511	3	US-08-464-052-2

28	44	28.0	511	4	US-08-461-002-2	Sequence 2, Appli
29	44	28.0	511	4	US-08-689-411-2	Sequence 2, Appli
30	44	28.0	511	5	PCT-US94-09863-2	Sequence 2, Appli
31	44	28.0	527	3	US-08-907-229-2	Sequence 2, Appli
32	43.5	27.7	602	3	US-08-446-100-1	Sequence 1, Appli
33	43.5	27.7	602	3	US-08-446-100-2	Sequence 2, Appli
34	43.5	27.7	602	3	US-08-446-100-3	Sequence 3, Appli
35	43.5	27.7	602	3	US-08-446-100-4	Sequence 4, Appli
36	43.5	27.7	602	3	US-08-446-100-5	Sequence 5, Appli
37	43.5	27.7	602	3	US-08-446-100-6	Sequence 6, Appli
38	43.5	27.7	602	3	US-08-446-100-7	Sequence 7, Appli
39	43.5	27.7	602	3	US-08-446-100-8	Sequence 8, Appli
40	43.5	27.7	602	3	US-08-446-100-11	Sequence 11, Appli
41	43.5	27.7	602	3	US-08-446-100-12	Sequence 12, Appli
42	43.5	27.7	602	3	US-08-446-100-13	Sequence 13, Appli
43	43.5	27.7	602	3	US-08-446-100-14	Sequence 14, Appli
44	43.5	27.7	602	3	US-08-446-100-15	Sequence 15, Appli
45	43.5	27.7	602	3	US-08-446-100-16	Sequence 16, Appli
46	43.5	27.7	602	3	US-08-446-100-17	Sequence 17, Appli
47	43.5	27.7	602	3	US-08-446-100-18	Sequence 18, Appli
48	43.5	27.7	602	3	US-08-446-100-24	Sequence 24, Appli
49	43.5	27.7	602	4	US-09-334-489-3	Sequence 3, Appli
50	43.5	27.7	602	4	US-09-334-489-4	Sequence 4, Appli
51	43.5	27.7	602	6	5215909-11	Patent No. 5215909
52	43	27.4	3066	4	US-08-952-127-12	Sequence 12, Appli
53	42.5	27.1	251	3	US-09-082-737-5	Sequence 5, Appli
54	42.5	27.1	273	4	US-09-060-410-18	Sequence 18, Appli
55	42.5	27.1	694	3	US-08-559-397A-31	Sequence 31, Appli
56	42.5	27.1	1183	2	US-08-447-031A-2	Sequence 2, Appli
57	42	26.8	450	2	US-08-955-138-12	Sequence 12, Appli
58	41.5	26.4	572	6	5200183-5	Patent No. 5200183
59	41	26.1	116	3	US-08-946-026-30	Sequence 30, Appli
60	41	26.1	244	2	US-08-553-633A-2	Sequence 2, Appli
61	41	26.1	244	2	US-08-737-825-2	Sequence 2, Appli
62	41	26.1	268	2	US-07-857-224B-20	Sequence 20, Appli
63	41	26.1	268	2	US-07-857-224B-21	Sequence 21, Appli
64	41	26.1	274	2	US-08-867-030B-15	Sequence 15, Appli
65	41	26.1	274	5	PCT-US95-06119-15	Sequence 15, Appli
66	41	26.1	294	4	US-09-230-637-31	Sequence 31, Appli
67	41	26.1	387	1	US-08-713-828-3	Sequence 3, Appli
68	41	26.1	387	2	US-08-919-627-3	Sequence 3, Appli
69	41	26.1	387	2	US-09-096-245-3	Sequence 3, Appli
70	41	26.1	388	1	US-08-713-828-4	Sequence 4, Appli
71	41	26.1	388	1	US-08-713-828-5	Sequence 5, Appli
72	41	26.1	388	2	US-08-919-627-4	Sequence 4, Appli
73	41	26.1	388	2	US-08-919-627-5	Sequence 5, Appli
74	41	26.1	388	2	US-09-096-245-4	Sequence 4, Appli
75	41	26.1	388	2	US-09-096-245-5	Sequence 5, Appli
76	41	26.1	436	4	US-09-150-213-4	Sequence 4, Appli
77	41	26.1	471	3	US-09-118-324-5	Sequence 5, Appli
78	41	26.1	471	4	US-09-286-691-12	Sequence 12, Appli
79	41	26.1	471	4	US-09-687-147-12	Sequence 12, Appli
80	41	26.1	492	1	US-07-794-393-4	Sequence 4, Appli
81	41	26.1	492	1	US-08-001-711-4	Sequence 4, Appli
82	41	26.1	643	2	US-08-245-511-47	Sequence 47, Appli
83	41	26.1	643	2	US-08-600-993A-47	Sequence 47, Appli
84	40.5	25.8	109	4	US-09-419-459-6	Sequence 6, Appli
85	40.5	25.8	268	2	US-08-852-743-3	Sequence 3, Appli
86	40.5	25.8	268	3	US-09-185-370-3	Sequence 3, Appli
87	40.5	25.8	390	4	US-09-419-459-4	Sequence 4, Appli
88	40.5	25.8	394	4	US-09-419-459-10	Sequence 10, Appli
89	40.5	25.8	465	2	US-08-114-555A-2	Sequence 2, Appli
90	40.5	25.8	465	3	US-08-559-397A-2	Sequence 2, Appli
91	40.5	25.8	506	1	US-08-369-780-2	Sequence 2, Appli
92	40.5	25.8	506	1	US-08-475-682-2	Sequence 2, Appli
93	40.5	25.8	506	1	US-08-780-833-2	Sequence 2, Appli
94	40.5	25.8	506	1	US-08-636-036-2	Sequence 2, Appli
95	40.5	25.8	506	3	US-08-918-509-2	Sequence 2, Appli
96	40.5	25.8	506	3	US-09-108-262-2	Sequence 2, Appli
97	40.5	25.8	515	4	US-08-961-083-22	Sequence 22, Appli
98	40.5	25.8	524	2	US-08-615-942A-2	Sequence 2, Appli
99	40.5	25.8	544	2	US-08-935-760-2	Sequence 2, Appli
100	40.5	25.8	544	3	US-08-559-397A-19	Sequence 19, Appli

## ALIGNMENTS

RESULT 1  
US-08-695-987-4  
; Sequence 4, Application US/08695987  
; Patent No. 6005090  
; GENERAL INFORMATION:  
; APPLICANT: DODGE, Christopher V.  
; APPLICANT: LEE, Adrian  
; APPLICANT: RADCLIFF, Fiona J.  
; APPLICANT: HAZELL, Stuart L.  
; TITLE OF INVENTION: TREATMENT AND PREVENTION OF HELIOBACTER  
; TITLE OF INVENTION: INFECTION  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/695,987  
; FILING DATE: 15-AUG-1996  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 17227/128  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 505 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-695-987-4

Query Match 100.0%; Score 157; DB 3; Length 505;  
Best Local Similarity 100.0%; Pred. No. 5.4e-17;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MNKDVQQTAFGAPVWDDNNVITAGPRG 29  
|||||  
Db 1 MNKDVQQTAFGAPVWDDNNVITAGPRG 29

RESULT 2  
US-657-868B-4  
; Sequence 4, Application US/08657868B  
; Patent No. 6080556  
; GENERAL INFORMATION:  
; APPLICANT: SUGIYAMA, Tosihiro  
; APPLICANT: KAWABATA, Tomohisa  
; APPLICANT: HIRAYASU, Kazunari  
; APPLICANT: TANAKA, Takumi  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION  
; TITLE OF INVENTION: AND USE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP

; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/657,868B  
; FILING DATE: 31-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 136465  
; FILING DATE: 02-JUN-1995  
; APPLICATION NUMBER: 83512  
; FILING DATE: 05-APR-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Resnick, David S.  
; REGISTRATION NUMBER: 34,235  
; REFERENCE/DOCKET NUMBER: 46522  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 505 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-657-868B-4

Query Match 100.0%; Score 157; DB 3; Length 505;  
Best Local Similarity 100.0%; Pred. No. 5.4e-17;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MNKDVQQTAFGAPVWDDNNVITAGPRG 29  
|||||  
Db 1 MNKDVQQTAFGAPVWDDNNVITAGPRG 29

RESULT 3  
US-08-657-868B-5  
; Sequence 5, Application US/08657868B  
; Patent No. 6080556  
; GENERAL INFORMATION:  
; APPLICANT: SUGIYAMA, Tosihiro  
; APPLICANT: KAWABATA, Tomohisa  
; APPLICANT: HIRAYASU, Kazunari  
; APPLICANT: TANAKA, Takumi  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION  
; TITLE OF INVENTION: AND USE  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/657,868B

; FILING DATE: 31-MAY-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 136465  
 ; FILING DATE: 02-JUN-1995  
 ; APPLICATION NUMBER: 83512  
 ; FILING DATE: 05-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Resnick, David S  
 ; REGISTRATION NUMBER: 34,235  
 ; REFERENCE/DOCKET NUMBER: 46522  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-6440  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; FRAGMENT TYPE: internal  
 ; US-08-657-868B-5

Query Match 100.0%; Score 157; DB 3; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-17;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNKDVKQTTFAGAPVWDNNVITAGPRG 29  
 Db 1 MVNKDVKQTTFAGAPVWDNNVITAGPRG 29

RESULT 4  
 US-08-657-868B-6  
 ; Sequence 6, Application US/08657868B  
 ; Patent No. 6080556  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUGIYAMA, Toshiro  
 ; APPLICANT: KAWABATA, Tomohisa  
 ; APPLICANT: HIRAYASU, Kazunari  
 ; APPLICANT: TANAKA, Takumi  
 ; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION  
 ; TITLE OF INVENTION: AND USE  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/657.868B  
 ; FILING DATE: 31-MAY-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 136465  
 ; FILING DATE: 02-JUN-1995  
 ; APPLICATION NUMBER: 83512  
 ; FILING DATE: 05-APR-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Resnick, David S  
 ; REGISTRATION NUMBER: 34,235  
 ; REFERENCE/DOCKET NUMBER: 46522  
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-6440  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; FRAGMENT TYPE: internal  
 ; US-08-657-868B-6

Query Match 100.0%; Score 157; DB 3; Length 505;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-17;  
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVNKDVKQTTFAGAPVWDNNVITAGPRG 29  
 Db 1 MVNKDVKQTTFAGAPVWDNNVITAGPRG 29

RESULT 5  
 US-08-695-987-2  
 ; Sequence 2, Application US/08695987  
 ; Patent No. 6005090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DOIDGE, Christopher V.  
 ; APPLICANT: LEE, Adrian  
 ; APPLICANT: RADCLIFF, Fiona J.  
 ; APPLICANT: HAZELL, Stuart L.  
 ; TITLE OF INVENTION: TREATMENT AND PREVENTION OF HELIOBACTER  
 ; TITLE OF INVENTION: INFECTION  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Foley & Lardner  
 ; STREET: 3000 K Street, N.W., Suite 500  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20007-5109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/695.987  
 ; FILING DATE: 15-AUG-1996  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: BENT, Stephen A.  
 ; REGISTRATION NUMBER: 29,768  
 ; REFERENCE/DOCKET NUMBER: 17227/128  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202)672-5300  
 ; TELEFAX: (202)672-5399  
 ; TELEX: 904136  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-695-987-2

Query Match 97.5%; Score 153; DB 3; Length 505;  
 Best Local Similarity 96.6%; Pred. No. 2.4e-16;  
 Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MVNKDVKQTTFAGAPVWDNNVITAGPRG 29



```

: COUNTRY: USA
: ZIP: 94306
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS
: SOFTWARE: Patent In Release

```

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/365,486A  
; FILING DATE: 23-DEC-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 8255-0018  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 527 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-365-486A-26

Query Match 35.7%; Score 56; DB 2; Length 527;  
Best Local Similarity 52.4%; Pred. No. 0.75;  
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAGAPVWDDNNVITAGPRG 29  
|| | | | | : : || | | |  
Db 28 TTGGNPIGDKLNTMTAGSRG 48

RESULT 12  
US-08-880-342-26  
; Sequence 26, Application US/08880342  
; Patent No. 6218179  
; GENERAL INFORMATION:  
; APPLICANT: Webster, Keith A.  
; APPLICANT: Bishopric, Nanette H.  
; APPLICANT: Murphy, Brian  
; APPLICANT: Laderoute, Keith R.  
; APPLICANT: Green, Christopher J.  
; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated  
; TITLE OF INVENTION: Therapeutic Constructs  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/880,342  
; FILING DATE: 23-JUN-1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/IB95/00996  
; FILING DATE: 13-NOV-1995

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/365,486  
; FILING DATE: 23-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 8255-0018.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 527 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-880-342-26

Query Match 35.7%; Score 56; DB 4; Length 527;  
Best Local Similarity 52.4%; Pred. No. 0.75;  
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 9 TTAGAPVWDDNNVITAGPRG 29  
|| | | | | : : || | | |  
Db 28 TTGGNPIGDKLNTMTAGSRG 48

RESULT 13  
US-07-846-181-5  
; Sequence 5, Application US/07846181  
; Patent No. 5360732  
; GENERAL INFORMATION:  
; APPLICANT: BERKA, RANDY M  
; APPLICANT: FOWLER, TIMOTHY  
; APPLICANT: REY, MICHAEL W  
; TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER  
; TITLE OF INVENTION: CATALASE-R  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
; STREET: 180 KIMBALL WAY  
; CITY: SOUTH SAN FRANCISCO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94080

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/846,181  
; FILING DATE: 19920304  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: HORN Ms, MARGARET A  
; REGISTRATION NUMBER: 33401  
; REFERENCE/DOCKET NUMBER: GC204-US1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-742-7536  
; TELEFAX: 415-742-7217  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 730 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-846-181-5

Query Match 33.4%; Score 52.5; DB 1; Length 730;  
Best Local Similarity 52.4%; Pred. No. 4;  
Matches 11; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 9 TTAGAPVWDDNNVITAGPRG 29  
|| | | | | : : || | | |  
Db 59 TTDFGTPI-SDQTSILKAGPRG 78

RESULT 14  
US-07-845-989-5  
; Sequence 5, Application US/07845989  
; Patent No. 5360901  
; GENERAL INFORMATION:

APPLICANT: BERKA, RANDY M  
APPLICANT: FOWLER, TIMOTHY  
APPLICANT: REY, MICHAEL W  
TITLE OF INVENTION: PRODUCTION OF ASPERGILLUS NIGER  
TITLE OF INVENTION: CATALASE-R  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENENCOR INTERNATIONAL, INC.  
STREET: 180 KIMBALL WAY  
CITY: SOUTH SAN FRANCISCO  
STATE: CA  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/845,989  
FILING DATE: 19920304  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: HORN MS, MARGARET A  
REGISTRATION NUMBER: 33401  
REFERENCE/DOCKET NUMBER: GC208-US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-742-7536  
TELEFAX: 415-742-7217  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 730 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-845-989-5

Query Match 33.4%; Score 52.5; DB 1; Length 730;  
Best Local Similarity 52.4%; Pred. No. 4;  
Matches 11; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

Qy 9 TTAFGAPVWDDNNVITAGPRG 29  
Db 59 TTDFCTPI-SDOTSLKAGPRG 78

RESULT 15  
US-08-201-697-4  
Sequence 4, Application US/08201697  
Patent No. 5705623  
GENERAL INFORMATION:  
APPLICANT: Wiggins, Roger C.  
APPLICANT: Thomas, Peedikayil E.  
TITLE OF INVENTION: Mammalian Glomerular Epithelial Protein  
TITLE OF INVENTION: 1  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/201,697  
FILING DATE: 25-FEB-1994  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UM 9783  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1188 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-201-697-4

Query Match 31.8%; Score 50; DB 1; Length 1188;  
Best Local Similarity 47.4%; Pred. No. 18;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 7 KOTTAFGAPVWDDNNVITA 25  
Db 25 KNATAFHVTVDDNNIVVS 43

RESULT 16  
US-09-061-768A-2  
Sequence 2, Application US/09061768A  
Patent No. 6204037  
GENERAL INFORMATION:  
APPLICANT: BRASH, ALAN R.  
APPLICANT: BOEGLIN, WILLIAM E.  
APPLICANT: JISAKA, MITSUO  
TITLE OF INVENTION: LIPOXYGENASE PROTEINS AND NUCLEIC ACIDS  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ARLES A. TAYLOR, JR.  
STREET: SUITE 1400, UNIVERSITY TOWER, 3100 TOWER BOULEVARD  
CITY: DURHAM  
STATE: NORTH CAROLINA  
COUNTRY: USA  
ZIP: 27707  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB storage  
COMPUTER: IBM PC/XT/AT compatible  
OPERATING SYSTEM: Windows 3.1  
SOFTWARE: WORD PERFECT 6.1 and ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/061,768A  
FILING DATE: APRIL 16, 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: NONE  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: ARLES A. TAYLOR, JR.  
REGISTRATION NUMBER: 39,395  
REFERENCE/DOCKET NUMBER: 1242/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 493-8000  
TELEFAX: (919) 419-0383  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 676 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
US-09-061-768A-2

Query Match 31.2%; Score 49; DB 4; Length 676;  
Best Local Similarity 52.6%; Pred. No. 13;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 11 AFGAPVDDNNVITAPRG 29  
|||| || :| ||  
Db 13 AFGAGTWDKVSIVGTRG 31

## RESULT 17

US-08-852-743-4  
; Sequence 4, Application US/08852743  
; Patent No. 5830699  
; GENERAL INFORMATION:  
; APPLICANT: Force, Thomas  
; APPLICANT: Kyriakis, John M.  
; APPLICANT: Pombo, Celia M.  
; APPLICANT: Bonventre, Joseph  
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/852,743  
; FILING DATE: 7-MAY-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/016,774  
; FILING DATE: 7-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00786/327001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-852-743-4

Query Match 28.3%; Score 44.5; DB 2; Length 271;  
Best Local Similarity 33.3%; Pred. No. 23;  
Matches 9; Conservative 6; Mismatches 9; Indels 3; Gaps 1;

QY 5 DVKQTTAFGAPVDDNNVITA---GPR 28  
:|:| | | | | :|:  
Db 153 NIKRTTWGTPYMWAPVSVSRKEYGPK 179

## RESULT 18

US-09-185-370-4  
; Sequence 4, Application US/09185370  
; Patent No. 6093560  
; GENERAL INFORMATION:  
; APPLICANT: Force, Thomas  
; APPLICANT: Kyriakis, John M.  
; APPLICANT: Pombo, Celia M.  
; APPLICANT: Bonventre, Joseph  
; TITLE OF INVENTION: SOK-1 AND METHODS OF USE

; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/185,370  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/852,743  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00786/327001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-185-370-4

Query Match 28.3%; Score 44.5; DB 3; Length 271;  
Best Local Similarity 33.3%; Pred. No. 23;  
Matches 9; Conservative 6; Mismatches 9; Indels 3; Gaps 1;

QY 5 DVKQTTAFGAPVDDNNVITA---GPR 28  
:|:| | | | | :|:  
Db 153 NIKRTTWGTPYMWAPVSVSRKEYGPK 179

## RESULT 19

US-08-689-411-8  
; Sequence 8, Application US/08689411  
; Patent No. 6224881  
; GENERAL INFORMATION:  
; APPLICANT: Riley M.D., Lee W.  
; APPLICANT: Chong, Pele  
; TITLE OF INVENTION: DNA MOLECULE FRAGMENTS ENCODING FOR  
; CELLULAR UPTAKE OF MYCOBACTERIUM TUBERCULOSIS AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/689,411  
; FILING DATE:  
; CLASSIFICATION: 536



ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/187  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 60 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-689-411-8

Query Match 28.0%; Score 44; DB 4; Length 60;  
Best Local Similarity 66.7%; Pred. No. 4.5;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VNKDVKQTTFG 13  
Db 1 VNADIKATTVFG 12

RESULT 20  
US-08-464-052-4  
Sequence 4, Application US/08464052  
Patent No. 6008201  
GENERAL INFORMATION:  
APPLICANT: Riley M.D., Lee W.  
TITLE OF INVENTION: DNA Molecule Encoding for Cellular  
Uptake of Mycobacterium Tuberculosis and Uses thereof  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,052  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/185 (D-1485B)  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 209 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: unknown  
US-08-464-052-4

Query Match 28.0%; Score 44; DB 3; Length 209;  
Best Local Similarity 66.7%; Pred. No. 20;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VNKDVKQTTFG 13

Db 1 VNADIKATTVFG 12

RESULT 21  
US-08-461-002-4  
Sequence 4, Application US/08461002  
Patent No. 6214543  
GENERAL INFORMATION:  
APPLICANT: Riley M.D., Lee W.  
TITLE OF INVENTION: DNA Molecule Encoding for Cellular  
Uptake of Mycobacterium Tuberculosis and Uses thereof  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,002  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldman, Michael L.  
REGISTRATION NUMBER: 30,727  
REFERENCE/DOCKET NUMBER: 19603/186 (D-1485B)  
TELEPHONE: (716) 263-1304  
TELEFAX: (716) 263-1600  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 209 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: unknown  
US-08-461-002-4

Query Match 28.0%; Score 44; DB 4; Length 209;  
Best Local Similarity 66.7%; Pred. No. 20;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VNKDVKQTTFG 13  
Db 1 VNADIKATTVFG 12

RESULT 22  
US-08-689-411-4  
Sequence 4, Application US/08689411  
Patent No. 6224881  
GENERAL INFORMATION:  
APPLICANT: Riley M.D., Lee W.  
APPLICANT: Chong, Pele  
TITLE OF INVENTION: DNA MOLECULE FRAGMENTS ENCODING FOR  
CELLULAR UPTAKE OF MYCOBACTERIUM TUBERCULOSIS AND USES  
THEREOF  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
STREET: Clinton Square, P.O. Box 1051  
CITY: Rochester  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 14603

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/689,411
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/187
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-689-411-4

Query Match      28.0%; Score 44; DB 4; Length 209;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 VNKDVKQTAFG 13
   |||||
DB 1 VNADIKATVFG 12

RESULT 23
US-07-794-393-2
; Sequence 2, Application US/07794393
; Patent No. 5236844
; GENERAL INFORMATION:
; APPLICANT: CHAMMON, PIERRE
; APPLICANT: BASSET, PAUL
; APPLICANT: BELLOCQ, JEAN-PIERRE
; TITLE OF INVENTION: ANALYTICAL MARKERS FOR MALIGNANT BREAST
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Ave. NW Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/794,393
; FILING DATE: 19911121
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9025326.1
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: GOLDSTEIN, JORGE A
; REGISTRATION NUMBER: 29,021
; REFERENCE/DOCKET NUMBER: 1383.0040000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 2:
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-794-393-2

Query Match      28.0%; Score 44; DB 1; Length 488;
Best Local Similarity 37.5%; Pred. No. 57;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVNKDVKQTTFAGAPVWDDNNVIT 24
   :|:|:|:|:|:|:|:|:|:|:|
DB 121 LVQEQVQRTWAEALKVWSDVTPLT 144

RESULT 24
US-08-001-711-2
; Sequence 2, Application US/08001711
; Patent No. 5484726
; GENERAL INFORMATION:
; APPLICANT: BASSET, PAUL
; APPLICANT: BELLOCQ, JEAN-PIERRE
; APPLICANT: CHAMMON, PIERRE
; TITLE OF INVENTION: ANALYTICAL MARKERS FOR MALIGNANT BREAST
; TITLE OF INVENTION: CANCER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Suite 300
; CITY: Washington
; STATE: D.C.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/001,711
; FILING DATE: 19930107
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/794,393
; FILING DATE: 11-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9025626.1
; FILING DATE: 21-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 1383.0040001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)466-0800
; TELEFAX: (202)833-8716
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 488 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-001-711-2

Query Match      28.0%; Score 44; DB 1; Length 488;
Best Local Similarity 37.5%; Pred. No. 57;
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 MVNKDVKQTTFAGAPVWDDNNVIT 24
   :|:|:|:|:|:|:|:|:|:|:|
DB 121 LVQEQVQRTWAEALKVWSDVTPLT 144

;
; INFORMATION FOR SEQ ID NO: 2:
```

RESULT 26  
US-08-448-489-11  
; Sequence 11, Application US/08448489  
; Patent No. 6184022  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: SATO, Hiroshi  
; APPLICANT: SHINAGAWA, Akira  
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR  
; FILE REFERENCE: 55-290P  
; CURRENT APPLICATION NUMBER: US/08/448,489

US-08-464-052-2

61 VNADTKATTVEG 72

RESULT 28  
US-08-461-002-2  
; Sequence 2, Application US/08461002  
; Patent No. 6214543  
; GENERAL INFORMATION:  
; APPLICANT: Riley M.D., Lee W.  
; TITLE OF INVENTION: DNA Molecule Encoding for Cellular  
; Uptake of Mycobacterium Tuberculosis and Uses thereof  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,002  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/186 (D-1485B)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 511 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-461-002-2

Query Match 28.0%; Score 44; DB 4; Length 511;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VNKDVKQTTFG 13  
|||:|||||  
Db 61 VNADIKATTVFG 72

RESULT 29  
US-08-689-411-2  
; Sequence 2, Application US/08689411  
; Patent No. 6224881  
; GENERAL INFORMATION:  
; APPLICANT: Riley M.D., Lee W.  
; APPLICANT: Chong, Pele  
; TITLE OF INVENTION: DNA MOLECULE FRAGMENTS ENCODING FOR  
; CELLULAR UPTAKE OF MYCOBACTERIUM TUBERCULOSIS AND USES  
; THEREOF  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/689,411  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman, Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/187  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1304  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 511 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-689-411-2

Query Match 28.0%; Score 44; DB 4; Length 511;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 VNKDVKQTTFG 13  
|||:|||||  
Db 61 VNADIKATTVFG 72

RESULT 30  
PCT-US94-09863-2  
; Sequence 2, Application PC/TUS9409863  
; GENERAL INFORMATION:  
; APPLICANT: Riley, Lee W.  
; TITLE OF INVENTION: DNA MOLECULE ENCODING FOR CELLULAR  
; UPTAKE OF MYCOBACTERIUM TUBERCULOSIS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Michael L. Goldman  
; STREET: Clinton Square, P.O. Box 1051  
; CITY: Rochester  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 14603  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09863  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldman Mr., Michael L.  
; REGISTRATION NUMBER: 30,727  
; REFERENCE/DOCKET NUMBER: 19603/180 (D-1485)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (716) 263-1000  
; TELEFAX: (716) 263-1600  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 511 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
PCT-US94-09863-2

Query Match 28.0%; Score 44; DB 5; Length 511;



Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

QY 2 VNKDVKQTAF---GAPVW---DDNNVIT 24  
|||| : ||| ||| : | : |||  
Db 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

## RESULT 34

US-08-446-100-3  
; Sequence 3, Application US/08446100  
; Patent No. 6001625  
; GENERAL INFORMATION:  
; APPLICANT: Broomfield, Clarence A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Lockridge, Oksana  
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hendricks and Assoc.  
; STREET: 9669 A Main Street, P.O. Box 2509  
; CITY: Fairfax  
; STATE: VA  
; COUNTRY: US  
; ZIP: 22031  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,100  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hendricks, Glenna  
; REGISTRATION NUMBER: 32,535  
; REFERENCE/DOCKET NUMBER: broomfield  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 425-4250  
; TELEFAX: (703) 425-2767  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 602 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ANTI-SENSE: YES  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: human esterases  
US-08-446-100-3

Query Match 27.7%; Score 43.5; DB 3; Length 602;  
Best Local Similarity 50.0%; Pred. No. 88;  
Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

QY 2 VNKDVKQTAF---GAPVW---DDNNVIT 24  
|||| : ||| ||| : | : |||  
Db 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

## RESULT 35

US-08-446-100-4  
; Sequence 4, Application US/08446100  
; Patent No. 6001625  
; GENERAL INFORMATION:  
; APPLICANT: Broomfield, Clarence A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Lockridge, Oksana  
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases

; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hendricks and Assoc.  
; STREET: 9669 A Main Street, P.O. Box 2509  
; CITY: Fairfax  
; STATE: VA  
; COUNTRY: US  
; ZIP: 22031  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,100  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hendricks, Glenna  
; REGISTRATION NUMBER: 32,535  
; REFERENCE/DOCKET NUMBER: broomfield  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 425-4250  
; TELEFAX: (703) 425-2767  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 602 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ANTI-SENSE: YES  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: human esterases  
US-08-446-100-4

Query Match 27.7%; Score 43.5; DB 3; Length 602;  
Best Local Similarity 50.0%; Pred. No. 88;  
Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

QY 2 VNKDVKQTAF---GAPVW---DDNNVIT 24  
|||| : ||| ||| : | : |||  
Db 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

## RESULT 36

US-08-446-100-5  
; Sequence 5, Application US/08446100  
; Patent No. 6001625  
; GENERAL INFORMATION:  
; APPLICANT: Broomfield, Clarence A  
; APPLICANT: Millard, Charles B  
; APPLICANT: Lockridge, Oksana  
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hendricks and Assoc.  
; STREET: 9669 A Main Street, P.O. Box 2509  
; CITY: Fairfax  
; STATE: VA  
; COUNTRY: US  
; ZIP: 22031  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,100  
; FILING DATE: 19-MAY-1995

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-5

```

Query Match 27.7%; Score 43.5; DB 3; Length 602;  
 Best Local Similarity 50.0%; Pred. No. 88;  
 Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

```

QY 2 VNKDVKQTAF---GAPVW--DDNNVIT 24
||||| : ||| ||| : |::||
DB 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

```

```

RESULT 37
US-08-446-100-6
; Sequence 6, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-2767
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES

```

```

; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-6

```

Query Match 27.7%; Score 43.5; DB 3; Length 602;  
 Best Local Similarity 50.0%; Pred. No. 88;  
 Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

```

QY 2 VNKDVKQTAF---GAPVW--DDNNVIT 24
||||| : ||| ||| : |::||
DB 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

```

```

RESULT 38
US-08-446-100-7
; Sequence 7, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
;
US-08-446-100-7

```

Query Match 27.7%; Score 43.5; DB 3; Length 602;  
 Best Local Similarity 50.0%; Pred. No. 88;  
 Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;

```

QY 2 VNKDVKQTAF---GAPVW--DDNNVIT 24
||||| : ||| ||| : |::||
DB 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374

```

```
RESULT 39
US-08-446-100-8
; Sequence 8, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELEPHONE: (703) 425-2767
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE: human esterases
; ORGANISM: human esterases
US-08-446-100-8
```

```
Query Match 27.7%; Score 43.5; DB 3; Length 602;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;
```

```
QY 2 VNKDVKQTAF---GAPVW--DDNNVIT 24
||||| : ||| : ||| : ||| : ||| : |||
Db 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374
```

```
RESULT 40
US-08-446-100-11
; Sequence 11, Application US/08446100
; Patent No. 6001625
; GENERAL INFORMATION:
; APPLICANT: Broomfield, Clarence A
; APPLICANT: Millard, Charles B
; APPLICANT: Lockridge, Oksana
; TITLE OF INVENTION: Site-Directed Mutagenesis of Esterases
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: 9669 A Main Street, P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: US
```

```
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,100
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELEPHONE: (703) 425-2767
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE: human esterases
; ORGANISM: human esterases
US-08-446-100-11
```

```
Query Match 27.7%; Score 43.5; DB 3; Length 602;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 14; Conservative 5; Mismatches 2; Indels 7; Gaps 3;
```

```
QY 2 VNKDVKQTAF---GAPVW--DDNNVIT 24
||||| : ||| : ||| : ||| : ||| : |||
Db 349 VNKD--EGTAFLVYGAPGFSKDNNSIIT 374
```

```
Search completed: August 22, 2002, 07:44:34
Job time: 93 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 22, 2002, 07:46:12 ; Search time 26.88 seconds  
(without alignments)  
103.668 Million cell updates/sec

Title: US-09-488-737-2

Perfect score: 157

Sequence: 1 MVNKDVKQTAFGAPVMDNNVITAGPRG 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	157	100.0	505	2	C64629 catalase (EC 1.11.
2	157	100.0	505	2	F71885 catalase (EC 1.11.
3	93	59.2	484	2	A58663 catalase (EC 1.11.
4	88	56.1	482	2	S60757 catalase (EC 1.11.
5	86	54.8	480	2	AB0148 catalase (EC 1.11.
6	83.5	53.2	491	1	CSR2 catalase (EC 1.11.
7	82	52.2	496	2	S37379 catalase (EC 1.11.
8	78.5	50.0	494	2	S62897 catalase (EC 1.11.
9	78	49.7	482	2	B83113 catalase PA4236 [1
10	76	48.4	492	2	S10395 catalase (EC 1.11.
11	76	48.4	492	2	S10770 catalase (EC 1.11.
12	75	47.8	492	2	T05779 catalase (EC 1.11.
13	74	47.1	492	2	S62696 catalase (EC 1.11.
14	74	47.1	507	2	G89908 Catalase [Imported
15	73	46.5	480	2	JC7672 catalase (EC 1.11.
16	73	46.5	492	2	T10902 catalase (EC 1.11.
17	72	45.9	493	2	T07911 catalase (EC 1.11.
18	71	45.2	483	2	JH0532 catalase (EC 1.11.
19	70	44.6	492	2	S48124 catalase (EC 1.11.
20	69	43.9	713	2	AC3220 catalase C [import
21	68	43.3	492	2	S17493 catalase (EC 1.11.
22	68	43.3	492	2	S46297 catalase (EC 1.11.
23	68	43.3	494	1	CSPM catalase (EC 1.11.
24	67	42.7	492	1	CSSY catalase (EC 1.11.
25	67	42.7	492	2	T16969 catalase (EC 1.11.
26	66	42.0	527	1	CSHU catalase (EC 1.11.
27	65	41.4	121	2	S65793 catalase (EC 1.11.
28	65	41.4	486	2	A57262 catalase (EC 1.11.
29	65	41.4	508	2	D64103 catalase (EC 1.11.

64	40.8	492	2	T06478 catalase (EC 1.11.
64	40.8	492	2	S52079 catalase (EC 1.11.
64	40.8	493	2	T12304 catalase (EC 1.11.
64	40.8	504	2	D81224 catalase (EC 1.11.
64	40.8	527	1	CSRT catalase (EC 1.11.
63	40.1	492	2	S71455 catalase (EC 1.11.
63	40.1	499	2	A55227 catalase (EC 1.11.
63	40.1	507	2	AD3621 catalase (EC 1.11.
62	39.5	482	2	S49465 catalase (EC 1.11.
62	39.5	705	2	B95843 probable catalase
60	38.2	349	2	C86862 hypothetical prote
60	38.2	492	2	S07124 catalase (EC 1.11.
41	38.2	492	2	T12300 catalase (EC 1.11.
42	37.6	492	2	T09756 catalase (EC 1.11.
43	37.6	492	2	T09754 catalase (EC 1.11.
44	37.6	492	2	T09754 catalase (EC 1.11.
45	36.9	485	2	S48650 catalase (EC 1.11.
58	36.9	488	2	T42038 catalase (EC 1.11.
57	36.3	249	2	T45217 rec7 protein - fis
48	36.3	339	2	T41126 meiotic recombinat
57	36.3	496	2	T37477 catalase (EC 1.11.
50	36.3	497	2	T42443 catalase (EC 1.11.
51	36.3	506	1	CSBO catalase (EC 1.11.
57	36.3	512	2	T27178 catalase (EC 1.11.
52	36.3	524	2	T27177 catalase (EC 1.11.
53	35.7	527	2	A36695 catalase (EC 1.11.
54	35.0	492	2	S46298 catalase (EC 1.11.
55	35.0	513	2	E83069 catalase PA4613 [1
56	34.4	489	2	N49388 catalase (EC 1.11.
57	34.4	492	2	JE0126 catalase (EC 1.11.
58	34.4	502	2	S40265 catalase (EC 1.11.
59	34.4	535	2	S65762 chitinase (EC 3.2.
60	34.4	567	2	C82183 catalase VC1585 [1
61	34.4	784	1	F64726 organic solvent to
62	34.4	784	2	C90636 organic solvent to
63	34.4	784	2	C85487 organic solvent to
64	34.4	784	2	AI0513 organic solvent to
54	33.8	169	1	S23478 probable benzoate
66	33.8	488	2	AH1422 catalase [imported
67	33.8	488	2	S37384 catalase (EC 1.11.
52	33.4	730	2	T26105 hypothetical prote
52	33.1	112	2	AI1796 catalase [imported
52	33.1	488	2	A44639 catalase (EC 1.11.
71	32.5	373	2	AF1957 hypothetical prote
72	32.5	488	2	A40367 catalase (EC 1.11.
73	31.8	252	2	C83877 transcription regu
50	31.8	327	2	F95395 probable LySR-fam
50	31.8	483	2	S37055 catalase (EC 1.11.
76	31.8	1188	1	A57064 protein-tyrosine-p
77	31.8	1216	2	S60613 protein-tyrosine-p
78	31.8	1216	2	S60613 protein-tyrosine-p
49	31.2	456	2	D70772 hypothetical prote
49	31.2	557	2	T39308 asparagine synthas
49	31.2	868	2	F89500 protein T27A10.6 [
48	30.6	474	2	S65763 chitinase (EC 3.2.
48	30.6	536	2	B75329 catalase (EC 1.11.
48	30.6	547	2	E69647 catalase (EC 1.11.
48	30.6	604	2	C87818 protein glh-1 [imp
48	30.6	604	2	T15132 ATP-dependent RNA
48	30.6	682	2	G86300 F19K9.13 protein
48	30.6	972	2	T50400 origin recognition
48	30.6	1226	2	JC7503 protein-tyrosine-p
48	30.6	1297	2	T39287 hypothetical prote
47	29.9	76	2	AD3194 hypothetical prote
47	29.9	187	2	T00427 ribosomal protein
47	29.9	187	2	H84916 60S ribosomal prot
47	29.9	307	2	A65170 hypothetical prote
47	29.9	331	2	A86713 cytochrome D ubiq
96	29.9	487	2	T34858 catalase (EC 1.11.
47	29.9	617	2	E64734 yach protein - Esc
47	29.9	707	2	AI3290 transketolase (EC
98	29.9	759	2	T42027 probable catalase
99	29.9	180	2	T44698 hypothetical prote
100	46.5	29.6		

## ALIGNMENTS

RESULT 1  
C64629  
catalase (EC 1.11.1.6) - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 04-Mar-2000  
C:Accession: C64629

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C. A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467

A:Accession: C64629

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-505 <TOM>

A:Cross-references: GB:AE000597; GB:AE000511; NID:g2314007; PIDN:AAD07923.1; PID:g231401

C:Superfamily: catalase

C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase

F:56,95,129/Active site: His, Ser, Asn #status predicted

F:339/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 100.0%; Score 157; DB 2; Length 505;  
Best Local Similarity 100.0%; Pred. No. 2e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNKKVKQTAFGAPVWDNNVITAGPRG 29

|||||

DB 1 MVNKKVKQTAFGAPVWDNNVITAGPRG 29

RESULT 2

F71885  
catalase (EC 1.11.1.6) - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 04-Mar-2000

C:Accession: F71885

R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;

Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path

A:Reference number: A71800; MUID:99120557

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-505 <ARN>

A:Cross-references: GB:AE001510; GB:AE001439; NID:g4155371; PIDN:AAD06391.1; PID:g415538

A:Experimental source: strain J99

C:Genetics:

A:Gene: kata

C:Superfamily: catalase

C:Keywords: heme; iron; metalloprotein; oxidoreductase

F:56,95,129/Active site: His, Ser, Asn #status predicted

F:339/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 100.0%; Score 157; DB 2; Length 505;  
Best Local Similarity 100.0%; Pred. No. 2e-15;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVNKKVKQTAFGAPVWDNNVITAGPRG 29

|||||

DB 1 MVNKKVKQTAFGAPVWDNNVITAGPRG 29

RESULT 3

A58663

catalase (EC 1.11.1.6) [validated] - Proteus mirabilis

C:Species: Proteus mirabilis

A:Variety: strain Pr, peroxide resistant

C:Date: 19-Nov-1997 #sequence\_revision 21-Nov-1997 #text\_change 15-Sep-2000

C:Accession: A58663; B58663

R:Buzy, A.; Bracchi, V.; Sterjiades, R.; Chroboczek, J.; Thibault, P.; Gagnon, J.; Jo

J. Protein Chem. 14, 59-72, 1995

A:Title: Complete amino acid sequence of Proteus mirabilis PR catalase. Occurrence of

A:Reference number: A58663; MUID:95305957

A:Accession: A58663

A:Molecule type: protein

A:Residues: 1-484 <BUZ1>

A:Accession: B58663

A:Molecule type: DNA

A:Residues: 1-305, 'AE', <BUZ2>

R:Gouet, P.; Jouve, H.M.; Dideberg, O.

submitted to the Brookhaven Protein Data Bank, June 1996

A:Reference number: A67899; PDB:2CAE

A:Contents: annotation; X-ray crystallography, 2.2 angstroms, residues 3-475

R:Gouet, P.; Jouve, H.M.; Hajdu, J.

submitted to the Brookhaven Protein Data Bank, June 1996

A:Reference number: A67900; PDB:2CAF

A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475

R:Note: compound 1

R:Gouet, P.; Jouve, H.M.; Hajdu, J.

submitted to the Brookhaven Protein Data Bank, June 1996

A:Reference number: A67901; PDB:2CAG

A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475

A:Note: compound II, dithiothreitol reduced compound I

R:Gouet, P.; Jouve, H.M.; Dideberg, O.

submitted to the Brookhaven Protein Data Bank, July 1996

A:Reference number: A67902; PDB:2CAH

A:Contents: annotation; X-ray crystallography, 2.7 angstroms, residues 3-475

A:Note: native Fe(III) with NADPH

R:Gouet, P.; Jouve, H.M.; Dideberg, O.

J. Mol. Biol. 249, 933-954, 1995

A:Title: Crystal structure of Proteus mirabilis PR catalase with and without bound NA

A:Reference number: A58664; MUID:95311317

A:Contents: annotation; X-ray crystallography, 2.2 angstroms

C:Complex: homotetramer

C:Function:

A:Description: catalyzes the conversion of two of molecules of hydrogen peroxide to t

A:Note: this enzyme has a tightly bound NADPH cofactor

C:Superfamily: catalase

C:Keywords: chromoprotein; heme; homotetramer; iron; metalloprotein; NADP; oxidoreduc

F:53/Modified site: methionine sulfone (Met) #status experimental

F:54,93,127/Active site: His, Ser, Asn #status predicted

F:337/Binding site: heme iron (Tyr) (axial ligand) #status experimental

Query Match 59.2%; Score 93; DB 2; Length 484;

Best Local Similarity 82.6%; Pred. No. 7.1e-06;

Matches 19; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 KOTTAFGAPVWDNNVITAGPRG 29

|||||

DB 5 KLTTAAGAPVWDNNVITAGPRG 27

RESULT 4

S60757

catalase (EC 1.11.1.6) - Bordetella pertussis

C:Species: Bordetella pertussis

C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 04-Mar-2000

C:Accession: S60757

R:DeShazer, D.; Wood, G.E.; Friedman, R.L.

Mol. Microbiol. 14, 123-130, 1994

A:Title: Molecular characterization of catalase from Bordetella pertussis: identifica

A:Reference number: S60757; MUID:95131725

A:Accession: S60757

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-482 <DES>

```

Query Match          52.2%;   Score 82;   DB 2;   Length 496;
Best Local Similarity 66.7%;   Pred. No. 0.00032;
Matches 14;   Conservative 2;   Mismatches 5;   Indels 0;   Gaps 0;

Qy      9  TTAFGAPVWDDNNVITAGPRG 29
      || |||||:|:| :| ||||
Db      20  TTNAGAPVWNDNEALTVGPRG 40

RESULT      8
S62697
catalase (EC 1.11.1.6) isoenzyme 2 - barley
C:Species: Hordeum vulgare (barley)
C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 04-Mar-2000
C:Accession: S62697
P:Skadsen, R.W.; Schulze-Lefert, P.; Herbst, J.M.
Plant Mol. Biol. 29, 1005-1014, 1995
A:Title: Molecular cloning, characterization and expression analysis of two catalase
A:Reference number: S62696; MUID:96145511
A:Accession: S62697
A>Status: preliminary; nucleic acid sequence not shown

```

```

Qy      9  TTAFGAPVWDDNNNVITAGPRG 29
      || |||||:::| ||||
Db     18  TTNSGAPVWNNSSSLTVGPRG 38

RESULT  11
SI0770
catalase (EC 1.11.1.6) - upland cotton
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 04-Mar-2000
C:Accession: SI0770
E:Ni, W.; Turley, R.B.; Trelease, R.N.
Biochim. Biophys. Acta 1049, 219-222, 1990
A:Title: Characterization of a cDNA encoding cottonseed catalase.
A:Reference number: SI0770; MUID:90304227
A:Accession: SI0770
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-492 <NIW>
A:Cross-references: GB:X52135; NID:g18483; PIDN:CAA36380.1; PID:g18484
C:Superfamily: catalase
C:Keywords: heme; iron; metalloprotein; oxidoreductase
F:65,104,138/Active site: His, Ser, Asn #status predicted
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match          48.4%; Score 76; DB 2; Length 492;
Best Local Similarity 61.9%; Pred. No. 0.0025;
Matches 13; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy      9  TTAFGAPVWDDNNNVITAGPRG 29
      || |||||:::| ||||
Db     18  TTNSGAPVWNNSSSLTVGPRG 38

RESULT  12
T05779
catalase (EC 1.11.1.6) - Arabidopsis thaliana
N:Alternate names: protein M4E13.140

```

C; Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 04-Mar-2000  
C; Accession: T05779; S18972  
R; Bevan, M.; Purnelle, B.; Boutry, M.; Goffeau, A.; Hoheisel, J.; Mewes, H.W.; Mayer, submitted to the Protein Sequence Database, April 1998  
A; Reference number: Z15451  
A; Accession: T05779  
A; Molecule type: DNA  
A; Residues: 1-492 <BEV>  
A; Cross-references: EMBL:AL022023  
A; Experimental source: cultivar Columbia; BAC clone M4E13  
R; Chevalier, C.; Yamaguchi, J.; McCourt, P.  
submitted to the EMBL Data Library, January 1992  
A; Description: Isolation and characterization of a catalase cDNA from Arabidopsis thaliana  
A; Reference number: S18972  
A; Accession: S18972  
A; Molecule type: mRNA  
A; Residues: 1-38; 'L', 40-108; 'A', 110-153; 'T', 155-242; 'L', 244-492 <CHE>  
A; Cross-references: EMBL:X64271; NID:g16214; PIDN:CAA45564.1; PID:g16215  
C; Genetics:  
A; Map position: 4  
A; Introns: 5/3; 38/1; 130/3; 389/3; 419/3; 442/2; 473/3  
C; Superfamily: catalase  
C; Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F; 65, 104, 138/Active site: His, Ser, Asn #status predicted  
F; 348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match	47.88;	Score 75;	DB 2;	Length 492;
Best Local Similarity	61.9%;	Pred. No. 0.0036;		
Matches 13;	Conservative 4;	Mismatches 0;	Gaps 0;	Indels 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29  
 || |||||:|:| |||||  
 Db 18 TTNSGAPVWNNSSMTVGPRG 38

RESULT 13  
 S62696  
 catalase (EC 1.11.1.6) isoenzyme 1 - barley  
 C:Species: Hordeum vulgare (barley)  
 C>Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 04-Mar-2000  
 C:Accession: S62696  
 R:Skadsen, R.W.; Schulze-Lefert, P.; Herbst, J.M.  
 Plant Mol. Biol. 29, 1005-1014, 1995  
 A:Title: Molecular cloning, characterization and expression analysis of two catalase iso  
 A:Reference number: S62696; MUID:96145511  
 A:Accession: S62696  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-492 <SKA>  
 A:Cross-references: EMBL:U20777; NID:g684945; PIDN:AAA96947.1; PID:g684946  
 A:Note: the authors did not translate the codon for residue 115  
 C:Superfamily: catalase  
 C:Keywords: heme; iron; metalloprotein; oxidoreductase  
 F:65,104,138/Active site: His, Ser, Asn #status predicted  
 F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 47.1%; Score 74; DB 2; Length 492;  
 Best Local Similarity 61.9%; Pred. No. 0.005;  
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29  
 || |||||:|:| |||||  
 Db 18 TTNSGAPVWNNALITVGRG 38

RESULT 14  
 G89908  
 Catalase [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: G89908  
 R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Babg, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C:Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146  
 A:Accession: G89908  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-507 <KUR>  
 A:Cross-references: GB:BA000018; PID:g13701132; PIDN:BAR42427.1; GSPDB:GN00149  
 A:Experimental source: strain N315  
 C:Genetics:  
 A:Gene: KatA  
 C:Superfamily: catalase

Query Match 47.1%; Score 74; DB 2; Length 507;  
 Best Local Similarity 55.2%; Pred. No. 0.0052;  
 Matches 16; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 1 MYNKDVKOTTAAGAPVWDDNNVITAGPRG 29  
 || |||||:|:| |||||  
 Db 3 MSQDDKLTGVFGHPVSDRENSMTAGPRG 31

RESULT 15  
 JC7672  
 catalase (EC 1.11.1.6) - Desulfovibrio vulgaris  
 C:Species: Desulfovibrio vulgaris  
 C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C:Accession: JC7672  
 R:Kitamura, M.; Nakanishi, T.; Kojima, S.; Kumagai, I.; Inoue, H.  
 J. Biochem. 129, 357-364, 2001  
 A:Title: Cloning and expression of the catalase gene from the anaerobic bacterium Des  
 A:Reference number: JC7672; MUID:21125589; PMID:11226874  
 A:Accession: JC7672  
 A:Molecule type: DNA  
 A:Residues: 1-480 <KIT>  
 A:Cross-references: DDBJ:AB020341  
 C:Comment: this enzyme, one of the hemoproteins, located in the cytoplasm, functions  
 C:Genetics:  
 A:Gene: kat  
 C:Superfamily: catalase  
 C:Keywords: Oxidoreductase

Query Match 46.5%; Score 73; DB 2; Length 480;  
 Best Local Similarity 65.2%; Pred. No. 0.0069;  
 Matches 15; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 7 KQTAFGAPVWDDNNVITAGPRG 29  
 || |||||:|:| |||||  
 Db 5 KLTNAGAPVPDQNANTAGPRG 27

RESULT 16  
 T10902  
 catalase (EC 1.11.1.6) - mung bean  
 C:Species: Vigna radiata (mung bean)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C:Accession: T10902  
 R:Mori, H.; Imaseki, H.  
 Plant Physiol. 102, 691-692, 1993  
 A:Title: CDNA for catalase from etioloated mung bean (Vigna radiata) hypocotyls.  
 A:Reference number: Z17208; MUID:94151449  
 A:Accession: T10902  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-492 <MOR>  
 A:Cross-references: EMBL:D13557; NID:g218321; PIDN:BAA02755.1; PID:g218322  
 A:Experimental source: clone VRCAT1  
 C:Function:  
 A:Description: catalyzes the conversion of two molecules of hydrogen peroxide to two  
 C:Superfamily: catalase  
 C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
 F:65,104,138/Active site: His, Ser, Asn #status predicted  
 F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 46.5%; Score 73; DB 2; Length 492;  
 Best Local Similarity 61.9%; Pred. No. 0.0071;  
 Matches 13; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAAGAPVWDDNNVITAGPRG 29  
 || |||||:|:| |||||  
 Db 18 TTNSGAPVWNNNSLTAVGTRG 38

RESULT 17  
 T07911  
 catalase (EC 1.11.1.6) - Chlamydomonas reinhardtii  
 C:Species: Chlamydomonas reinhardtii  
 C>Date: 21-May-1999 #sequence\_revision 21-May-1999 #text\_change 20-Jun-2000  
 C:Accession: T07911; T08127  
 R:Kim, J.Y.; Lee, K.O.; Lee, S.H.  
 submitted to the EMBL data Library, August 1997  
 A:Reference number: Z16209  
 A:Accession: T07911  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-493 <KIM>  
 A:Cross-references: EMBL:AF016902; NID:g2388690; PIDN:AAB70006.1; PID:g2388691  
 A:Experimental source: strain CC-125 wild type mt+ 137c



A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-713 <KUP>  
A:Cross-references: GB:AE008687; PIDN:AAL46177.1; PID:g17743948; GSPDB:GN00188  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: catE  
A:Genome: plasmid

Query Match 43.9%; Score 69; DB 2; Length 713;  
Best Local Similarity 66.7%; Pred. No. 0.043;  
Matches 14; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
||| ||||| :| :| |||||  
DB 40 TTAQGAPVADDDNSIRIGPRG 60

RESULT 21  
SI7493  
catalase (EC 1.11.1.6) - upland cotton  
C:Species: Gossypium hirsutum (upland cotton)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 04-Mar-2000  
C:Accession: SI7493  
R:Ni, W.; Trelease, R.N.  
Arch. Biochem. Biophys. 289, 237-243, 1991  
A:Title: Two genes encode the two subunits of cottonseed catalase.  
A:Reference number: SI7493; MUID:91378551  
A:Accession: SI7493  
A:Molecule type: mRNA  
A:Residues: 1-492 <NIW>  
A:Cross-references: EMBL:X56675; NID:g18487; PIDN:CAA39998.1; PID:g18488  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:65,104,138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 43.3%; Score 68; DB 2; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.04;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
||| ||||| :| :| |||||  
DB 18 TTNSGAPVWNNSSLTGVRG 38

RESULT 22  
S46297  
catalase (EC 1.11.1.6) catI - castor bean  
C:Species: Ricinus communis (castor bean)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Jun-2000  
C:Accession: S46297  
R:Suzuki, M.; Ario, T.; Hattori, T.; Nakamura, K.; Asahi, T.  
Plant Mol. Biol. 25, 507-516, 1994  
A:Title: Isolation and characterization of two tightly linked catalase genes from castor  
A:Reference number: S46297; MUID:94325474  
A:Accession: S46297  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-492 <SUZ>  
A:Cross-references: GB:D21161; NID:g532341; PIDN:BAA04697.1; PID:g643601  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:65,104,138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 43.3%; Score 68; DB 2; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.04;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
||| ||||| :| :| |||||  
DB 18 TTNSGAPVWNNSSLTGVRG 38

RESULT 23  
CSPM  
catalase (EC 1.11.1.6) - garden pea  
C:Species: Pisum sativum (garden pea)  
C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 03-Mar-2000  
C:Accession: SI8346; SI5559  
R:Isin, S.H.; Allen, R.D.  
Plant Mol. Biol. 17, 1263-1265, 1991  
A:Title: Isolation and characterization of a pea catalase cDNA.  
A:Reference number: SI8346; MUID:92032793  
A:Accession: SI8346  
A:Molecule type: mRNA  
A:Residues: 1-494 <ISI>  
A:Cross-references: EMBL:X60169; NID:g20676; PIDN:CAA42736.1; PID:g20677  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase; peroxisome; te  
F:65,104,138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 43.3%; Score 68; DB 1; Length 494;  
Best Local Similarity 57.1%; Pred. No. 0.04;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
||| ||||| :| :| |||||  
DB 18 TTNSGAPVWNNSSLTGVRG 38

RESULT 24  
CSSY  
catalase (EC 1.11.1.6) - soybean  
C:Species: Glycine max (soybean)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 03-Mar-2000  
C:Accession: S20999  
R:Allen, R.  
submitted to the EMBL Data Library, May 1992  
A:Reference number: S20999  
A:Accession: S20999  
A:Molecule type: DNA  
A:Residues: 1-492 <ALL>  
A:Cross-references: EMBL:Z12021; NID:g18559; PIDN:CAA78056.1; PID:g18560  
C:Genetics: 5/3; 38/1; 389/3; 419/3; 442/2; 473/3  
C:Introns:  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:65,104,138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 42.7%; Score 67; DB 1; Length 492;  
Best Local Similarity 52.4%; Pred. No. 0.056;  
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
||| ||||| :| :| |||||  
DB 18 TTNSGAPIWNNSSLTGVRG 38

RESULT 25  
T16969  
catalase (EC 1.11.1.6) 3 - curled-leaved tobacco  
C:Species: Nicotiana plumbaginifolia (curled-leaved tobacco)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T16969  
R:Willekens, H.; Villarroel, R.; van Montagu, M.; Inze, D.; van Camp, W.  
FEBS Lett. 352, 79-83, 1994  
A:Title: Molecular identification of catalases from Nicotiana plumbaginifolia (L.).

A:Reference number: S48650; MUID:95010681  
A:Accession: T16969  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-492 <WIL>  
A:Cross-references: EMBL:Z36977; NID:g536786; PIDN:CAA85426.1; PID:g536787  
A:Experimental source: flower  
C:Genetics:  
A:Gene: cat3  
C:Function:  
A:Description: catalyzes the conversion of two molecules of hydrogenperoxide to two mole  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:65,104,138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 42.7%; Score 67; DB 2; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.056;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||:::|:| | ||  
Db 18 TTNSCAPVWNNSSMTVGTGR 38  
|| |||||:::|:| | ||

RESULT 26  
CSHU  
catalase (EC 1.1.1.6) precursor [validated] - human  
A:Alternate names: hydrogen peroxidase; hydrogen peroxide oxidoreductase  
C:Species: Homo sapiens (man)  
C>Date: 17-Mar-1987 #sequence\_revision 18-Aug-1995 #text\_change 08-Dec-2000  
C:Accession: A23646; A23651; A00501; I55606  
R:Quan, F.; Korneluk, R.G.; Tropak, M.B.; Gravel, R.A.  
Nucleic Acids Res. 14, 5321-5335, 1986  
A:Title: Isolation and characterization of the human catalase gene.  
A:Reference number: A23646; MUID:86286546  
A:Accession: A23646  
A:Molecule type: DNA  
A:Residues: 1-527 <QUA>  
A:Cross-references: GB:X04085; NID:g29682; PIDN:CAA27721.1; PID:g1228085  
R:Bell, G.I.; Najarian, R.C.; Mullenbach, G.T.; Hallewell, R.A.  
Nucleic Acids Res. 14, 5561-5562, 1986  
A:Title: cDNA sequence coding for human kidney catalase.  
A:Reference number: A23651; MUID:86286565  
A:Accession: A23651  
A:Molecule type: mRNA  
A:Residues: 1-527 <BEL>  
A:Cross-references: GB:X04076; NID:g29720; PIDN:CAA27717.1; PID:g29721  
R:Korneluk, R.G.; Quan, F.; Lewis, W.H.; Guise, K.S.; Willard, H.F.; Holmes, M.T.; Gravel  
J. Biol. Chem. 259, 13819-13823, 1984  
A:Title: Isolation of human fibroblast catalase cDNA clones. Sequence of clones derived  
A:Reference number: A00501; MUID:85054813  
A:Accession: A00501  
A:Molecule type: mRNA  
A:Residues: 1-527 <KOR>  
A:Cross-references: GB:K02400; NID:g940256; PIDN:AAB59522.1; PID:g179950  
A:Experimental source: fibroblast  
A:Note: the carboxyl-terminal may be subject to tissue specific processing  
R:Yoo, J.H.; Erzurum, S.C.; Hay, J.G.; Lemarchand, P.; Crystal, R.G.  
J. Clin. Invest. 93, 297-302, 1994  
A:Title: Vulnerability of the human airway epithelium to hyperoxia. Constitutive express  
A:Reference number: I55606; MUID:94110338  
A:Accession: I55606  
A:Status: translation not shown; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-22 <RES>  
A:Cross-references: GB:U13609; NID:g388711; PIDN:AAA15651.1; PID:g388712  
R:Schröder, W.A.; Shelton, J.R.; Shelton, J.B.; Apell, G.; Evans, L.; Bonaventura, J.;  
Arch. Biochem. Biophys. 214, 422-424, 1982  
A:Title: The partial amino acid sequence of human erythrocyte catalase.  
A:Reference number: A44717; MUID:82205126  
A:Contents: annotation; sequence of peptides differing in composition from bovine

C:Comment: Catalase occurs in almost all aerobically respiring organisms and serves t  
C:Genetics:  
A:Gene: GDB:CAT  
A:Cross-references: GDB:119049; OMIM:115500  
A:Map position: 11p13-11p13  
A:Introns: 22/3; 80/1; 117/1; 160/3; 195/3; 237/3; 301/3; 352/3; 399/1; 442/3; 478/3;  
C:Complex: homotetramer  
C:Function:  
A:Description: catalyzes the conversion of two of molecules of hydrogen peroxide to t  
C:Superfamily: catalase  
C:Keywords: blocked amino end; chromoprotein; heme; homotetramer; iron; metalloprotei  
F:2-527/Product: catalase #status experimental <MAT>  
F:2-519/Product: catalase, peroxisomal form #status predicted <MAT2>  
F:517-519/Region: peroxisome/glyoxysome location signal (S-[RKH]-L) motif  
F:2/Modified site: blocked amino end (Ala) (in mature form) (probably acetylated) #st  
F:75,114,148/Active site: His, Ser, Asn #status predicted  
F:358/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 42.0%; Score 66; DB 1; Length 527;  
Best Local Similarity 66.7%; Pred. No. 0.086;  
Matches 14; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||:::|:| | ||  
Db 28 TTGAGNPVGDKLNVTGPRG 48  
|| |||||:::|:| | ||

RESULT 27  
S65793  
catalase (EC 1.1.1.6) - guinea pig (fragment)  
C:Species: Cavia porcellus (guinea pig)  
C>Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 06-Jun-1997  
C:Accession: S65793  
R:Yuan, H.T.; Bingle, C.D.; Kelly, F.J.  
Biochim. Biophys. Acta 1305, 163-171, 1996  
A:Title: Differential patterns of antioxidant enzyme mRNA expression in guinea pig lu  
A:Reference number: S65793; MUID:96180320  
A:Accession: S65793  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-121 <YUA>  
A:Cross-references: EMBL:U99841  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; oxidoreductase

Query Match 41.4%; Score 65; DB 2; Length 121;  
Best Local Similarity 61.9%; Pred. No. 0.022;  
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||:::|:| | ||  
Db 19 TTAGNPVGDKLNVTGPRG 39  
|| |||||:::|:| | ||

RESULT 28  
A57262  
catalase (EC 1.1.1.6) - Bacteroides fragilis  
C:Species: Bacteroides fragilis  
C>Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 04-Mar-2000  
C:Accession: A57262  
R:Rocha, E.R.; Smith, C.J.  
J. Bacteriol. 177, 3111-3119, 1995  
A:Title: Biochemical and genetic analyses of a catalase from the anaerobic bacterium  
A:Reference number: A57262; MUID:95286491  
A:Accession: A57262  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-486 <ROC>  
A:Cross-references: GB:U18676; NID:g841191; PIDN:AAC43384.1; PID:g841192  
C:Genetics:  
A:Gene: katB



C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:54.93.127/Active site: His, Ser, Asn #status predicted  
F:337/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 41.4%; Score 65; DB 2; Length 486;  
Best Local Similarity 55.2%; Pred. No. 0.11;  
Matches 16; Conservative 2; Mismatches 9; Indels 2; Gaps 1;

QY 1 MVNKDVKQTAFGAPVWDDNNVITAGPRG 29  
DB 1 MENK--KLTAANGRPDIADNQNQSQTAGPRG 27

RESULT 29

D64103  
catalase (EC 1.11.1.6) - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 04-Mar-2000  
C:Accession: D64103  
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Smali, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A.; Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: D64103  
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-508 <TIQR>  
A:Cross-references: GB:U32774; GB:L42023; NID:g1573942; PIDN:AAC22587.1; PID:g1573949; T  
C:Comment: This predicted catalase is unusual in having Thr-102 replace the catalytical  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:63.102.136/Active site: His, Thr, Asn #status predicted  
F:346/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 41.4%; Score 65; DB 2; Length 508;  
Best Local Similarity 61.9%; Pred. No. 0.12;  
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
DB 16 TMNGAPVADNQNQSLTAGPRG 36

RESULT 30

T06478  
catalase (EC 1.11.1.6) - wheat  
C:Species: Triticum aestivum (common wheat)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000  
C:Accession: T06478  
R: Saruyama, H.  
submitted to the EMBL Data Library, July 1996  
A:Reference number: 215707  
A:Accession: T06478  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-492 <SAR>  
A:Cross-references: EMBL:D86327; PIDN:BAAL3068.1  
A:Experimental source: seedling  
C:Genetics:  
A:Gene: catl  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:65.104.138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.8%; Score 64; DB 2; Length 492;

Best Local Similarity 52.4%; Pred. No. 0.16;  
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
DB 18 STNSGAPVWNNDSLTGSRG 38

RESULT 31

SS2079  
catalase (EC 1.11.1.6) - common sunflower  
C:Species: Helianthus annuus (common sunflower)  
C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 04-Mar-2000  
C:Accession: SS2079  
R: Kleff, S.; Trelease, R.N.; Eising, R.  
Biochim. Biophys. Acta 1224, 463-466, 1994  
A:Title: Nucleotide and deduced amino acid sequence of a putative higher molecular we  
A:Reference number: SS2079; MUID:95101741  
A:Accession: SS2079  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-492 <KLE>  
A:Cross-references: EMBL:I28740; NID:g453528; PIDN:AAA69866.1; PID:g453529  
A:Note: the authors translated the codon TTC for residue 103 as Leu  
C:Superfamily: catalase  
C:Keywords: heme; iron; metalloprotein; oxidoreductase  
F:65.104.138/Active site: His, Ser, Asn #status predicted  
F:348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.8%; Score 64; DB 2; Length 492;  
Best Local Similarity 57.1%; Pred. No. 0.16;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
DB 18 TTNSGAPVWNNNSLTGSRG 38

RESULT 32

TI2304  
catalase (EC 1.11.1.6), root - common ice plant  
C:Species: Mesembryanthemum crystallinum (common ice plant)  
C:Date: 23-Jul-1999 #sequence\_revision 23-Jul-1999 #text\_change 04-Mar-2000  
C:Accession: TI2304  
R:Michalowski, C.B.; Ougley-Landreau, F.; Bohnert, H.J.  
submitted to the EMBL Data Library, June 1998  
A:Description: Different isoforms of catalase are expressed in leaves and roots of th  
A:Reference number: 217487  
A:Accession: TI2304  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-493 <MIC>  
A:Cross-references: EMBL:AF069320; NID:g3202033; PID:g3202034  
C:Superfamily: catalase  
C:Keywords: heme; iron; metalloprotein; oxidoreductase  
F:65.104.138/Active site: His, Ser, Asn #status predicted  
F:349/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.8%; Score 64; DB 2; Length 493;  
Best Local Similarity 52.4%; Pred. No. 0.16;  
Matches 11; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 9 TTAFCGAPVWDDNNVITAGPRG 29  
DB 18 TTKTGPVWDDSLTVGARG 38

RESULT 33

DB1224  
catalase (EC 1.11.1.6) NMA0050 [similarity] - Neisseria meningitidis (strain MC58 ser  
C:Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001  
C;Accession: D81224; C81996  
R;Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.  
Science 287, 1809-1815, 2000  
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V  
A;Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A;Reference number: A81000; MUID:20175755  
A;Accession: D81224  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-504 <TET>  
A;Cross-references: GB:AE002379; GB:AE002098; NID:g7225435; PIDN:AAF40672.1; PID:g722543  
A;Experimental source: serogroup B, strain MC58  
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
Nature 404, 502-506, 2000  
A;Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* 22491.  
A;Reference number: A81775; MUID:20222556  
A;Accession: C81996  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-504 <PAR>  
A;Cross-references: GB:AL162759; GB:AL157959; NID:g7378778; PIDN:CAB83367.1; PID:g737862  
A;Experimental source: serogroup A, strain 22491  
C;Genetics:  
A;Gene: kata; NMB0216; NMA0050  
C;Superfamily: catalase  
C;Keywords: heme; iron; metalloprotein; oxidoreductase  
F;59,99,132/Active site: His, Thr, Asn #status predicted  
F;342/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.8% Score 64; DB 2; Length 504;  
Best Local Similarity 70.6% Pred. No. 0.16;  
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 13 GAPVDDNNVITAGPRG 29  
||||| : | : |||||  
Db 16 GAPVADNQSITAGPRG 32

RESULT 34  
CSRT  
catalase (EC 1.11.1.6) - rat  
N;Alternate names: hydrogen peroxidase; hydrogen peroxide oxidoreductase  
C;Species: *Rattus norvegicus* (Norway rat)  
C;Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 03-Mar-2000  
C;Accession: JU0065; A25965; I52202  
R;Nakashima, H.; Yamamoto, M.; Goto, K.; Osumi, T.; Hashimoto, T.; Endo, H.  
Gene 79, 279-288, 1989  
A;Title: Isolation and characterization of the rat catalase-encoding gene.  
A;Reference number: JU0065; MUID:90006757  
A;Accession: JU0065  
A;Molecule type: DNA  
A;Residues: 1-527 <NAK>  
A;Cross-references: GB:M25680; GB:M23742; NID:g203333; PIDN:AAB42378.1; PID:g203335  
R;Furuta, S.; Hayashi, H.; Hijikata, M.; Miyazawa, S.; Osumi, T.; Hashimoto, T.  
Proc. Natl. Acad. Sci. U.S.A. 83, 313-317, 1986  
A;Title: Complete nucleotide sequence of cDNA and deduced amino acid sequence of rat liv  
A;Reference number: A25965; MUID:86094381  
A;Accession: A25965  
A;Molecule type: mRNA  
A;Residues: 1-527 <FUR>  
A;Cross-references: GB:M11670; NID:g203344; PIDN:AAA40884.1; PID:g203345  
A;Experimental source: liver  
R;Osumi, T.; Ozasa, H.; Miyazawa, S.; Hashimoto, T.  
Biochem. Biophys. Res. Commun. 122, 831-837, 1984  
A;Title: Molecular cloning of cDNA for rat liver catalase.  
A;Reference number: I52202; MUID:84280086  
A;Accession: I52202  
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA  
A;Residues: 399-433,'N',435-527 <RES>  
A;Cross-references: GB:K01929; NID:g203346; PIDN:AAA40885.1; PID:g203347  
C;Genetics:  
A;Gene: cat  
A;Introns: 22/3; 80/1; 117/1; 160/3; 195/3; 237/3; 301/3; 352/3; 399/1; 442/3; 478/3;  
C;Complex: homotetramer  
C;Function:  
A;Description: catalyzes the conversion of two of molecules of hydrogen peroxide to t  
C;Superfamily: catalase  
C;Keywords: chromoprotein; heme; homotetramer; iron; metalloprotein; oxidoreductase;  
F;75,114,148/Active site: His, Ser, Asn #status predicted  
F;358/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.8% Score 64; DB 1; Length 527;  
Best Local Similarity 57.1% Pred. No. 0.17;  
Matches 12; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 9 TTAFGAPVDDNNVITAGPRG 29  
|| | | | : |||||  
Db 28 TTGGGNPIGDKLIMTAGPRG 48

RESULT 35  
S71455  
catalase (EC 1.11.1.6) 2 - maize  
C;Species: *Zea mays* (maize)  
C;Date: 12-Feb-1998 #sequence\_revision 13-Mar-1998 #text\_change 04-Mar-2000  
C;Accession: S71455; S18819; A36062  
R;Guan, L.; Polidoros, A.N.; Scandalios, J.G.  
Plant Mol. Biol. 30, 913-924, 1996  
A;Title: Isolation, characterization and expression of the maize Cat2 catalase gene.  
A;Reference number: S71455; MUID:96270370  
A;Accession: S71455  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-492 <GUA>  
A;Cross-references: EMBL:X54819; NID:g22233; PIDN:CAA38588.1; PID:g22234  
A;Note: due to an additional T the nucleic acid sequence submitted to the EMBL Data L  
A;Note: the authors translated the codon AAC for residue 28 as Met, GCC for residue 2  
R;Guan, L.; Ruzsa, S.; Skadsen, R.W.; Scandalios, J.G.  
Plant Physiol. 96, 1379-1381, 1991  
A;Title: Comparison of the cat2 complementary DNA sequences of a normal catalase acti  
A;Reference number: S18819  
A;Accession: S18819  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-240,'RC',244-296,'T',298-492 <GUW>  
A;Cross-references: EMBL:X54819; NID:g22233; PIDN:CAA38588.1; PID:g22234  
R;Bethards, L.A.; Skadsen, R.W.; Scandalios, J.G.  
Proc. Natl. Acad. Sci. U.S.A. 87, 6927, 1990  
A;Reference number: A36062; MUID:90370897  
A;Contents: erratum  
A;Accession: A36062  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-123,'N',125-240,'RC',244-296,'R',298-492 <BET>  
A;Cross-references: GB:J02976  
A;Note: the authors translated the codon AAC for residue 124 as Lys  
C;Genetics:  
A;Gene: cat2  
A;Introns: 5/3; 38/1; 419/3; 442/2; 473/3  
C;Superfamily: catalase  
C;Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F;64,103,137/Active site: His, Ser, Asn #status predicted  
F;348/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.1% Score 63; DB 2; Length 492;  
Best Local Similarity 52.4% Pred. No. 0.22;  
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 18 TTNSGAPVWDDSLVTGARG 38

## RESULT 36

A55227  
catalase (EC 1.11.1.6) - Brucella abortus  
C:Species: Brucella abortus  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 04-Mar-2000  
C:Accession: A55227  
R:Sha, Z.; Stabel, T.J.; Mayfield, J.E.  
J. Bacteriol. 176, 7375-7377, 1994  
A:Title: Brucella abortus catalase is a periplasmic protein lacking a standard signal sequence  
A:Reference number: A55227; MUID:95050323  
A:Accession: A55227  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-499 <SHA>  
A:Cross-references: GB:U11439; NID:g508707; PIDN:AAA64655.1; PID:g508708  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; homotetramer; iron; metalloprotein; oxidoreductase; periplasmic; protein  
F:2-499/Product: catalase #status experimental <MA>  
F:55,93,127/Active site: His, Ser, Asn #status predicted  
F:337/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 40.1%; Score 63; DB 2; Length 499;  
Best Local Similarity 57.1%; Pred. No. 0.23;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 8 TTNSGAPVWDDSLVTGARG 28

## RESULT 37

AD3621  
catalase (EC 1.11.1.6) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C:Accession: AD3621  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, A.; Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AD3621  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-507 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AAL54135.1; PID:g17985098; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10893  
A:Map position: II  
C:Superfamily: catalase  
C:Keywords: oxidoreductase

Query Match 40.1%; Score 63; DB 2; Length 507;  
Best Local Similarity 57.1%; Pred. No. 0.23;  
Matches 12; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 16 TTNSGAPVWDDSLVTGARG 36

## RESULT 38

S49465  
catalase (EC 1.11.1.6) - nematode (Onchocerca volvulus)  
C:Species: Onchocerca volvulus

C:Date: 01-Feb-1995 #sequence\_revision 10-Nov-1995 #text\_change 04-Mar-2000  
C:Accession: S49465  
R:Henkle-Duehrsen, K.J.  
submitted to the EMBL Data Library, October 1994  
A:Reference number: S49452  
A:Accession: S49465  
A:Molecule type: mRNA  
A:Residues: 1-482 <HEN>  
A:Cross-references: EMBL:X82176; NID:g558624; PIDN:CAA57666.1; PID:g558625  
C:Superfamily: catalase  
C:Keywords: chromoprotein; heme; iron; metalloprotein; oxidoreductase  
F:55,94,128/Active site: His, Ser, Asn #status predicted  
F:338/Binding site: heme iron (Tyr) (axial ligand) #status predicted

Query Match 39.5%; Score 62; DB 2; Length 482;  
Best Local Similarity 61.9%; Pred. No. 0.31;  
Matches 13; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 8 TTNSGAPVWDDSLVTGARG 28

## RESULT 39

B95843  
probable catalase (EC 1.11.1.6) [imported] - Sinorhizobium meliloti (strain 1021) mag  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 14-Sep-2001  
C:Accession: B95843  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herzig, A.; Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1.683-Kb pSymB megaplasmid from the N2-fixing eucalyptus root nodule bacterium Sinorhizobium meliloti  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: B95843  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-705 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC48410.1; PID:g15139882; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymB  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubli, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.; Hyman, R.W.; Jones, T.L.; Hyman, R.W.; Jones, T.L.; Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A.; Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: catC; Smb20007  
A:Genome: plasmid  
C:Keywords: oxidoreductase

Query Match 39.5%; Score 62; DB 2; Length 705;  
Best Local Similarity 61.9%; Pred. No. 0.48;  
Matches 13; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 9 TTAFCAPVWDDNNVITAGPRG 29  
|| |||||::: || ||  
Db 41 TTAQGPVADDDNSLRGARG 61

## RESULT 40

C86662  
hypothetical protein ycjI [imported] - Lactococcus lactis subsp. lactis (strain IL140)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: C86662  
R:Boletín, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehlen, T.; et al. The complete genome sequence of the lactic acid bacterium Lactococcus lactis  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis

A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: C86662  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-349 <STO>  
A;Cross-references: GB:AE005176; PID:gl2723163; PIDN:AAK04397.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: ycjI

Query Match 38.2%; Score 60; DB 2; Length 349;  
Best Local Similarity 56.5%; Pred. No. 0.42;  
Matches 13; Conservative 1; Mismatches 7; Indels 2; Gaps 1;  
  
Qy 7 KOTTAFGA--PVWDDNNVITAGP 27  
| | | | | : | | | | |  
Db 256 KWTTFGAKESIWDDTNVIGVTP 278

Search completed: August 22, 2002, 07:46:14  
Job time: 192 sec